

日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日 2003年11月20日
Date of Application:

出願番号 特願2003-390239
Application Number:
[ST. 10/C]: [JP 2003-390239]

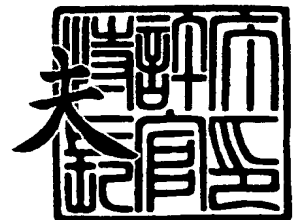
出願人 株式会社日立製作所
Applicant(s):

U.S. Appln. Filed 2-5-04
Inventor: Y. Watanabe et al
Mattingly Stanger & Malor
Docket NIT-409

2003年12月17日

特許庁長官
Commissioner,
Japan Patent Office

今井 康



出証番号 出証特2003-3104757

【書類名】 特許願
【整理番号】 NT03P0856
【提出日】 平成15年11月20日
【あて先】 特許庁長官 殿
【国際特許分類】 G06F 12/00
【発明者】
 【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所
 システム開発研究所内
 【氏名】 渡辺 恭男
【発明者】
 【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所
 システム開発研究所内
 【氏名】 山本 康友
【発明者】
 【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所
 システム開発研究所内
 【氏名】 江口 賢哲
【特許出願人】
 【識別番号】 000005108
 【氏名又は名称】 株式会社日立製作所
【代理人】
 【識別番号】 100068504
 【弁理士】
 【氏名又は名称】 小川 勝男
 【電話番号】 03-3537-1621
【選任した代理人】
 【識別番号】 100086656
 【弁理士】
 【氏名又は名称】 田中 恭助
 【電話番号】 03-3537-1621
【選任した代理人】
 【識別番号】 100094352
 【弁理士】
 【氏名又は名称】 佐々木 孝
 【電話番号】 03-3537-1621
【手数料の表示】
 【予納台帳番号】 081423
 【納付金額】 21,000円
【提出物件の目録】
 【物件名】 特許請求の範囲 1
 【物件名】 明細書 1
 【物件名】 図面 1
 【物件名】 要約書 1

【書類名】 特許請求の範囲**【請求項 1】**

ホストとのインターフェースであるポートと、キャッシュメモリと、前記ポート、前記キャッシュメモリ及び共有メモリと接続線により接続される制御装置と、該制御装置に接続されるディスク装置とを備えたストレージ装置であって、

前記ホストに提供する論理デバイスに対する第一の優先度情報を受け付ける管理端末を接続し、

前記制御装置は、前記管理端末から受け付けた第一の優先度情報に従って、優先度の高い論理デバイスは優先度の低い論理デバイスよりも多くの物理デバイスを割り当てておき、障害発生時には前記キャッシュメモリ内に格納されている論理デバイスのデータをこの論理デバイスに対応付けられている複数の物理デバイスに格納するように制御することを特徴とするストレージ装置。

【請求項 2】

前記第一の優先度情報に従って割り当てられた論理デバイスと複数の物理デバイスとの対応関係を示す管理テーブルが前記共有メモリに格納されていることを特徴とする請求項 1 記載のストレージ装置。

【請求項 3】

さらに、前記管理端末は、業務の優先度を示す第二の優先度情報を受け付けるように構成し、

前記制御装置は、障害発生時には前記管理端末から受け付けた第二の優先度情報に従って、ストレージ内のプライオリティーキューの優先度が高いジョブをデキューし、該ジョブを実行するように制御することを特徴とする請求項 1 または 2 記載のストレージ装置。

【請求項 4】

前記管理端末で受け付けた第二の優先度情報からなる管理テーブルを前記共有メモリに格納されていることを特徴とする請求項 3 記載のストレージ装置。

【請求項 5】

ホストとのインターフェースであるポートと、キャッシュメモリと、前記ポート、前記キャッシュメモリ及び共有メモリと接続線により接続される制御装置と、該制御装置に接続されるディスク装置とを備えたストレージ装置であって、

業務の優先度を示す第二の優先度情報を受け付ける管理端末を接続し、

前記制御装置は、障害発生時には前記管理端末で受け付けた第二の優先度情報に従って、ストレージ内のプライオリティーキューの優先度が高いジョブをデキューし、該ジョブを実行するように制御することを特徴とするストレージ装置。

【請求項 6】

前記第二の優先度情報からなる管理テーブルを前記共有メモリに格納して構成することを特徴とする請求項 5 記載のストレージ装置。

【請求項 7】

前記第二の優先度情報が前記ホスト優先度情報であることを特徴とする請求項 5 または 6 記載のストレージ装置。

【請求項 8】

前記制御装置は、障害発生時に対象論理デバイスのデータを消失した際、該消失したデータに対応する論理デバイスを識別するための情報を前記管理テーブルに登録することを特徴とする請求項 1 乃至 7 の何れかに記載のストレージ装置。

【請求項 9】

前記制御装置は、障害発生時に対象論理デバイスのデータを消失した際、該消失したデータに対応する論理デバイス及び論理デバイス内の位置を識別するための情報を表示装置に表示することを特徴とする請求項 1 乃至 6 の何れかに記載のストレージ装置。

【請求項 10】

ホストとのインターフェースであるポートと、キャッシュメモリと、前記ポート、前記キャッシュメモリ及び共有メモリと接続線により接続される制御装置と、該制御装置に接

続されるディスク装置とを備えたストレージ装置における制御方法であって、

前記ホストに提供する論理デバイスに対する第一の優先度情報を受け付ける第1のステップと、

該第1のステップにおいて受け付けた第一の優先度情報に従って、優先度の高い論理デバイスは優先度の低い論理デバイスよりも多くの物理デバイスを割り当てておき、障害発生時には前記キャッシュメモリ内に格納されている論理デバイスのデータをこの論理デバイスに対応付けられている複数の物理デバイスに格納する第2のステップとを有することを特徴とするストレージ装置における制御方法。

【請求項11】

更に、該第1のステップにおいて受け付けた第一の優先度情報に従って割り当てられた論理デバイスと複数の物理デバイスとの対応関係を示す管理テーブルを前記共有メモリに格納しておく第3のステップを有することを特徴とする請求項10記載のストレージ装置における制御方法。

【請求項12】

さらに、業務の優先度を示す第二の優先度情報を受け付ける第4のステップと、

障害発生時には前記第4のステップから受け付けた第二の優先度情報に従って、ストレージ内のプライオリティーキューの優先度が高いジョブをデキューし、該ジョブを実行するように制御する第5のステップとを有することを特徴とする請求項10または11記載のストレージ装置における制御方法。

【請求項13】

さらに、前記第4のステップで受け付けた第二の優先度情報からなる管理テーブルを前記共有メモリに格納しておく第6のステップを有することを特徴とする請求項12記載のストレージ装置における制御方法。

【請求項14】

ホストとのインターフェースであるポートと、キャッシュメモリと、前記ポート、前記キャッシュメモリ及び共有メモリと接続線により接続される制御装置と、該制御装置に接続されるディスク装置とを備えたストレージ装置における制御プログラムであって、

前記ホストに提供する論理デバイスに対する第一の優先度情報を受け付ける第1のステップと、

該第1のステップにおいて受け付けた第一の優先度情報に従って、優先度の高い論理デバイスは優先度の低い論理デバイスよりも多くの物理デバイスを割り当てておき、障害発生時には前記キャッシュメモリ内に格納されている論理デバイスのデータをこの論理デバイスに対応付けられている複数の物理デバイスに格納する第2のステップとを有することを特徴とするストレージ装置における制御プログラム。

【請求項15】

更に、該第1のステップにおいて受け付けた第一の優先度情報に従って割り当てられた論理デバイスと複数の物理デバイスとの対応関係を示す管理テーブルを前記共有メモリに格納しておく第3のステップを有することを特徴とする請求項14記載のストレージ装置における制御プログラム。

【請求項16】

さらに、業務の優先度を示す第二の優先度情報を受け付ける第4のステップと、

障害発生時には前記第4のステップから受け付けた第二の優先度情報に従って、ストレージ内のプライオリティーキューの優先度が高いジョブをデキューし、該ジョブを実行するように制御する第5のステップとを有することを特徴とする請求項14または15記載のストレージ装置における制御プログラム。

【請求項17】

さらに、前記第4のステップで受け付けた第二の優先度情報からなる管理テーブルを前記共有メモリに格納しておく第6のステップを有することを特徴とする請求項16記載のストレージ装置における制御プログラム。

【請求項18】

ホストとのインターフェースであるポートと、キャッシュメモリと、前記ポート、前記キャッシュメモリ及び共有メモリと接続線により接続される制御装置と、該制御装置に接続されるディスク装置とを備えたストレージ装置におけるジョブスケジューリング処理方法であって、

障害が発生している状態であるかを判定する第7のステップと、

該第7のステップで障害が発生していると判定された場合には、ジョブ管理テーブルに登録されているジョブが接続されている論理デバイスプライオリティーキュー、ホストプライオリティーキュー、業務プライオリティーキューの何れかのプライオリティーキューの優先度の高いジョブをデキューし、該デキューしたジョブを実行する第8のステップとを有することを特徴とするストレージ装置におけるジョブスケジューリング処理方法。

【請求項19】

前記第8のステップにおいて、前記プライオリティーキューの選択には、論理デバイスの優先度、ホスト優先度及び業務優先度の比率を用いることを特徴とする請求項18記載のストレージ装置におけるジョブスケジューリング処理方法。

【請求項20】

さらに、論理デバイス管理テーブルに登録された、キャッシュメモリ上のダーティデータがディスク装置に完全に退避しているか否かの障害発生時データ退避情報をチェックし、その結果完全に退避していないことを示す論理デバイスに関してアクセスパターン管理テーブルに登録されているダーティデータ量の現在値を参照して値がゼロであれば前記障害発生時データ退避情報を更新する第9のステップを有することを特徴とする請求項18または19記載のストレージ装置におけるジョブスケジューリング処理方法。

【請求項21】

ホストとのインターフェースであるポートと、キャッシュメモリと、前記ポート、前記キャッシュメモリ及び共有メモリと接続線により接続される制御装置と、該制御装置に接続されるディスク装置とを備えたストレージ装置におけるジョブスケジューリング処理プログラムであって、

障害が発生している状態であるかを判定する第7のステップと、

該第7のステップで障害が発生していると判定された場合には、ジョブ管理テーブルに登録されているジョブが接続されている論理デバイスプライオリティーキュー、ホストプライオリティーキュー、業務プライオリティーキューの何れかのプライオリティーキューの優先度の高いジョブをデキューし、該デキューしたジョブを実行する第8のステップとを有することを特徴とするストレージ装置におけるジョブスケジューリング処理プログラム。

【請求項22】

前記第8のステップにおいて、前記プライオリティーキューの選択には、論理デバイスの優先度、ホスト優先度及び業務優先度の比率を用いることを特徴とする請求項21記載のストレージ装置におけるジョブスケジューリング処理プログラム。

【請求項23】

ホストとのインターフェースであるポートと、キャッシュメモリと、前記ポート、前記キャッシュメモリ及び共有メモリと接続線により接続される制御装置と、該制御装置に接続されるディスク装置とを備えたストレージ装置における障害処理方法であって、

障害処理を終了するか否かを判定する第10のステップと、

該第10のステップで処理を終了しない場合、ストレージ装置内の各部位に障害が発生しているかをチェックする第11のステップと、

該第11のステップのチェック結果で障害が検出された場合には、キャッシュメモリ上にダーティデータを有する論理デバイスに関する障害発生時データ退避情報を設定する第12のステップと、

その後、各論理デバイスに関するキャッシュメモリ上のダーティデータを物理デバイスに反映させるために各論理デバイスのダーティデータに関するライトジョブのエンキュー処理を行う第13のステップと、

その後、対象ライトジョブの完了を待ち、この過程で予備電源不足若しくはキャッシュメモリの二重障害が検出されたときに、対象論理デバイスのデバイス状態を障害オフラインに設定する第14のステップとを有することを特徴とするストレージ装置における障害処理方法。

【請求項24】

ホストとのインターフェースであるポートと、キャッシュメモリと、前記ポート、前記キャッシュメモリ及び共有メモリと接続線により接続される制御装置と、該制御装置に接続されるディスク装置とを備えたストレージ装置における障害処理プログラムであって、

障害処理を終了するか否かを判定する第10のステップと、

該第10のステップで処理を終了しない場合、ストレージ装置内の各部位に障害が発生しているかをチェックする第11のステップと、

該第11のステップのチェック結果で障害が検出された場合には、キャッシュメモリ上にダーティデータを有する論理デバイスに関する障害発生時データ退避情報を設定する第12のステップと、

その後、各論理デバイスに関するキャッシュメモリ上のダーティデータを物理デバイスに反映させるために各論理デバイスのダーティデータに関するライトジョブのエンキュー処理を行う第13のステップと、

その後、対象ライトジョブの完了を待ち、この過程で予備電源不足若しくはキャッシュメモリの二重障害が検出されたときに、対象論理デバイスのデバイス状態を障害オフラインに設定する第14のステップとを有することを特徴とするストレージ装置における障害処理プログラム。

【書類名】 明細書

【発明の名称】 ストレージ装置、ストレージ装置における制御方法、ジョブスケジューリング処理方法及び障害処理方法並びにそれらのプログラム

【技術分野】**【0001】**

本発明は、接続線によりディスク装置と接続され、更に相互結合網によりキャッシュメモリおよび共有メモリと結合され、ホストとのインターフェースであるポートを介して前記ホストとの間で取り交わす入出力要求を処理する制御装置を有するストレージ装置（計算機システム）、ストレージ装置における制御方法、ジョブスケジューリング処理方法及び障害処理方法並びにそれらのプログラムに関し、特に、障害発生時のデータ優先度制御技術に関する。

【背景技術】**【0002】**

従来技術 1 としては、特開平 11-167521 号公報に記載された技術が知られている。この従来技術 1 には、コモンバス方式を採用することにより、システム構成（規模）に応じてホストアダプタ、記憶装置アダプタ等の各論理モジュールやキャッシュメモリ及びディスク装置を接続することでスケラブルなシステムを実現することができるようにすると共に、各論理モジュール、ディスク装置及びコモンバスの多重化により、縮退運転と各論理モジュール及び記憶媒体の活線挿抜対応とを可能とし、無停止で保守することができる記憶システムが記載されている。

【0003】

上記従来技術 1 のように、ホストと大容量のディスク装置との間に介在し、両者の間のデータ転送を制御するディスク制御装置の多くは、転送されるデータの一時記憶機構としてのキャッシュメモリを備えている。しかし、このキャッシュメモリは揮発性のメモリのために、電力供給が停止すると格納されているデータが消失する。また、ディスク制御装置におけるキャッシュメモリのハードウェア故障により、データが消失する可能性もある。このようなデータ消失を回避するために、二重化されたキャッシュメモリを有し、書き込みデータを二重化されたキャッシュメモリに格納するディスク制御装置が知られている。しかしながら、キャッシュメモリに二重障害が発生した場合には、データは消失することになる。

【0004】

さらに、上記従来技術 1 は、また、冗長化された電源を有し、いずれかの電源が故障しても、継続運転可能となっている。また、停電時にはディスク制御装置が有する予備電源によりキャッシュメモリ、共有メモリ上のデータの消失を回避することができる。電源故障および停電時にはキャッシュメモリ上のデータ保障のため、キャッシュメモリ上のデータデータをディスク装置に反映する。

【0005】

さらに、以後のホストからのライト要求に関しては、データ保障のために、ディスク装置にライトデータが書き込まれた後にホストに対してライト完了報告を行うライト処理である同期ライトを用いる。しかし、該同期ライトにはライトアфтаと比べてホストへの応答性能が低いという課題がある。

【0006】

また、従来技術 2 としては、特開 2002-334049 号公報に記載された技術が知られている。この従来技術 2 には、非同期リモートコピーを行う複数ホスト配下の記憶サブシステム内でのサイドファイル量の流入量制限をホストに対して設定した優先度に基づいて行い、重要度の低いデータがキャッシュメモリを圧迫するのを防ぐことが記載されている。

【0007】

また、従来技術 3 としては、特開 2003-6016 号公報に記載された技術が知られている。この従来技術 3 には、非同期リモートコピーにおいて、論理ボリュームグループ

にコピーの優先度を付け、優先度の高い論理ボリュームグループのコピーを優先することが記載されている。

【0008】

【特許文献1】特開平11-167521号公報

【0009】

【特許文献2】特開2002-334049号公報

【特許文献3】特開2003-6016号公報

【発明の開示】

【発明が解決しようとする課題】

【0010】

上記従来技術1においては、障害発生時に、キャッシュメモリ上のダーティデータが消失する恐れがある。また、電源障害に備えて予備電源を有し、キャッシュメモリ上のダーティデータの保全を行っている。また、キャッシュメモリの障害に備えて、キャッシュメモリおよびキャッシュメモリを管理するためのデータを保持する共有メモリを二重化している。

【0011】

しかしながら、上記従来技術1においては、停電が継続し予備電源が使用できなくなった場合、キャッシュメモリもしくは共有メモリに二重障害が発生した場合にはディスクに反映されていないデータはデータの重要度とは無関係に消失する課題を有していた。また、障害発生時にはデータ保障の観点から同期ライトを用いるために、ストレージの入出力処理全体としての性能低下が発生するという課題を有していた。

【0012】

また、上記従来技術2および3は、障害発生時（電源障害やキャッシュメモリの障害）に関しては考慮されていないものである。

【0013】

本発明の目的は、上記課題を解決すべく、障害発生時にキャッシュメモリ上のデータを速やかにディスク装置に退避し、優先度の高い重要なデータの消失を回避することができるストレージ装置、ストレージ装置における制御方法、ジョブスケジューリング処理方法及び障害処理方法並びにそれらのプログラムを提供することにある。

【課題を解決するための手段】

【0014】

ストレージ装置は、ホストとのインターフェースであるポートと、キャッシュメモリと、前記ポート、前記キャッシュメモリ及び共有メモリと接続線により接続される制御装置と、該制御装置に接続されるディスク装置とを備える。ストレージ装置はホストに提供する論理デバイスに対する第一の優先度情報を管理端末から受け付け、第一の優先度情報に従って、優先度の高い論理デバイスは優先度の低い論理デバイスよりも多くの物理デバイスを割り当てておき、障害発生時には前記キャッシュメモリ内に格納されている論理デバイスのデータをこの論理デバイスに対応付けられている複数の物理デバイスに格納するように制御する。

【発明の効果】

【0015】

本発明によれば、ストレージ装置に障害が発生した際に、ストレージ内のキャッシュメモリ上のダーティデータを速やかにディスク装置に反映することができ、優先度の高い重要なデータの消失を回避することができる。

【0016】

また、本発明によれば、ストレージ装置に障害が発生した際に、優先度の高い重要な業務に関する性能低下を極力回避することができる。

【発明を実施するための最良の形態】

【0017】

以下、本発明に係る実施の形態を図面を用いて説明する。

【0018】

本発明の実施の形態は、キャッシュメモリ112上にあつて、ディスク装置114に対して未反映なデータであるダーティデータを、ストレージ110に障害が発生した際に、速やかにディスク装置114に反映するために、論理デバイスの配置先の物理デバイスをあらかじめ最適にしておくものである。また、ストレージ110に障害が発生した際には、ストレージ110は、あらかじめストレージ管理者からの入力によって運用管理される管理端末140により論理デバイスに設定された優先度順にダーティデータをディスク装置114に反映する。さらに、ストレージ110に障害が発生した際には、ストレージ110は、あらかじめストレージ管理者からの入力によって運用管理される管理端末140により設定された優先度に基づきストレージ内のジョブをスケジューリングする。

【0019】

次に、図1から図26を参照して、本発明の実施の形態を説明する。

【0020】

図1は本発明に係る計算機システム（ストレージ装置）1の一実施の形態を説明するものである。計算機システム1は、ホスト100、ストレージ110及び管理端末140を備えて構成される。ホスト100はストレージ110に対して入出力要求を行う。ストレージ110はホスト100からの入出力要求に対して、ストレージ110が保持するキャッシュメモリ112を介して、ディスク装置114に読み書きを行う。管理端末140はストレージ管理者からの入力を受付け、ストレージ110の運用管理を行う。

【0021】

次に、ストレージ110の構成について説明する。ストレージ110は、チャンネルアダプタ120、ディスクアダプタ130、相互結合網（接続線から構成される。）111、キャッシュメモリ112、共有メモリ113、ディスク装置114、接続線115、電源116、予備電源（バッテリー）117及び電力線118を備えて構成される。そして、ディスクアダプタ130とディスク装置114との間は、1つのディスクアダプタあるいは1つの相互結合網において障害が発生した時にもディスク装置114が使用できるように1つのディスク装置に2つのディスクアダプタ130が異なる相互結合網111で接続されている。

【0022】

チャンネルアダプタ120は、ホスト100とキャッシュメモリ112との間のデータ転送を制御する。ディスクアダプタ130は、キャッシュメモリ112とディスク装置114との間のデータ転送を制御する。キャッシュメモリ112はホスト100から受信したデータ、あるいは、ディスク装置114から読み取ったデータを一時的に蓄えるメモリである。通常、ホスト100からのライト要求時には、ホスト100への応答性能の向上を目的として、キャッシュメモリ112にデータが書き込まれた時点で、ホスト100にライトの完了報告を行うライトアフト方式が用いられる。共有メモリ113は、全てのチャンネルアダプタ120とディスクアダプタ130とが共有するメモリである。

【0023】

ディスク装置114を用いたストレージ、特にRAID（Redundant Array of Independent Disks）などのディスクアレイを制御するストレージでは、実際に搭載した物理デバイスであるディスク装置114上に、ホスト100へ提供する論理デバイスの規定のデータ形式に従いデータを格納している。

【0024】

チャンネルアダプタ120、ディスクアダプタ130、キャッシュメモリ112、共有メモリ113、電源116及び予備電源117は、障害対応のため二重化されている。チャンネルアダプタ120は、ホスト100からポート121を介して入出力要求を受け、ホスト100とのデータの転送を制御する。ディスクアダプタ130は、ディスク装置114とのデータの転送を制御する。チャンネルアダプタ120、ディスクアダプタ130は共に、データの転送をキャッシュメモリ112を介して行う。電源116および予備電源117は、電力線118を介して、チャンネルアダプタ120、ディスクアダプタ130、キャ

ッシュメモリ112、共有メモリ113及びディスク装置114に電力を供給する。キャッシュメモリ112は、ホスト100からの書込みデータ、ディスク装置114からの読み込みデータを一時的に蓄えるメモリである。通常、ホスト100からのライト要求時には、ホスト100への応答性能の向上を目的として、キャッシュメモリ112にデータが書き込まれた時点で、ホスト100にライトの完了報告を行うライトアフタ方式が用いられる。さらに、キャッシュメモリ112は、一定量に分割して管理されており、分割したキャッシュメモリをセグメントと呼ぶ。各セグメントの状態は、後述するセグメント管理テーブル6で管理される。共有メモリ113は、全てのチャネルアダプタ120とディスクアダプタ130とが共有するメモリである。そして、共有メモリ113には、キャッシュメモリ112上のデータを制御するための情報、チャネルアダプタ120上の制御プロセッサで動作するジョブを制御するための情報、ディスクアダプタ130上の制御プロセッサで動作するジョブを制御するための情報を含む。即ち、共有メモリ113には、図2に示すように、論理デバイス管理テーブル2、LUパス管理テーブル3、物理デバイス管理テーブル4、スロット管理テーブル5、セグメント管理テーブル6、チャネルジョブ管理テーブル7、ディスクジョブ管理テーブル8、ホスト管理テーブル9、アクセスパターン管理テーブル10、業務管理テーブル11及びスケジューリング管理テーブル12等の情報が記憶されることになる。

【0025】

管理端末140は、PCなどのストレージ管理プログラム動作手段142～145と、ストレージ管理者との入出力手段141を有し、論理デバイスに関する属性設定など、I/F146、119を介しての、ストレージ110の保守運用に関するストレージ管理者からのインターフェースとなる。

【0026】

以上説明したように、ストレージ110は、ホスト100とのインターフェースであるポート121を有し、上記ホストとの間で取り交わす入出力要求を処理するチャネルアダプタ120及びディスクアダプタ130から構成される制御装置を有し、該制御装置120、130は接続線115によりディスク装置114と接続され、さらに相互結合網111によりキャッシュメモリ112及び共有メモリ113と結合され、ストレージ管理者より各種パラメータ（図14及び図15並びに図16、図17及び図18に示す論理デバイス定義処理及びLUパス定義処理並びに論理デバイス優先度定義処理、ホスト優先度定義処理及び業務優先度定義処理についての各種パラメータ）を受け付ける管理端末140とI/F119により接続し、該管理端末140で受け付けられた各種パラメータを管理テーブル2～11として格納する共有メモリ113で構成される。

【0027】

次に、論理デバイス管理テーブル2の一実施例について図3を用いて説明する。論理デバイス管理テーブル2は、論理デバイス番号201、サイズ202、デバイス状態203、物理デバイス番号204、接続ホスト番号205、ポート番号／ターゲットID/LUN206、論理デバイス優先度207、拡張論理デバイス情報208及び障害発生時データ退避情報209を含む。

【0028】

サイズ202には、当該論理デバイスの容量が格納される。例えば、図3では、論理デバイス番号1、2、…に関して、サイズに1GB、2GB、…を指定している。

デバイス状態203には、当該論理デバイスの状態を示す情報が設定される。状態としては、「オンライン」「オフライン」「未実装」「障害オフライン」が存在する。「オフライン」は、当該論理デバイスは定義され、正常に稼動しているが、LUパス未定義などでホスト100からのアクセスはできない状態にあることを示す。「未実装」は、当該論理デバイスが定義されておらずホスト100からのアクセスはできない状態にあることを示す。「障害オフライン」は、当該論理デバイスに障害が発生してホスト100からのアクセスができないことを示す。「オンライン」である論理デバイスに関して障害が検出された時点で、当該論理デバイスに関して「障害オフライン」が設定される。例えば、図3

では、論理デバイス番号 1、2、…に関して、デバイス状態に「オンライン」、「オンライン」、…を指定している。

【0029】

物理デバイス番号 204 には、当該論理デバイスが対応する物理デバイス番号が格納される。例えば、図 3 では、論理デバイス番号 1、2、…に関して、物理デバイス番号に 1、2、…を指定している。

接続ホスト番号 205 は、当該論理デバイスにアクセスが許可されているホスト 100 を識別するホスト番号である。例えば、図 3 では、論理デバイス番号 1、2、…に関して、接続ホスト番号に 1、2、…を指定している。

エントリ 206 のポート番号には、当該論理デバイスが複数のポート 121 のどのポート 121 に接続されているかを表す情報が設定される。各ポート 121 には、ストレージ 110 内で一意な番号が割り振られており、当該論理デバイスが LUN 定義されているポート 121 の番号が記録される。また、同エントリのターゲット ID と LUN は、論理デバイスを識別するための識別子である。ここでは、これらの識別子として、SCSI 上でホスト 100 からデバイスをアクセスする場合に用いられる SCSI-ID、LUN が用いられる。例えば、図 3 では、論理デバイス番号 1、2、…に関して、ポート ID に 1、2、…、ターゲット ID/LUN に 0/0、0/0、…を指定している。

論理デバイス優先度 207 は、ストレージ 110 での障害発生時におけるデータの優先度を示す情報であり、例えば、最高優先度の値 1 から最低優先度の値 5 までの数値を用い、図 3 では、論理デバイス番号 1、2、…に関して、論理デバイス優先度に 1、5、…を指定している。

拡張論理デバイス情報 208 は、複数の論理デバイスを連結し、ひとつの論理デバイス（拡張論理デバイスと呼ぶ）としてホスト 100 に提供するための情報であり、拡張論理デバイスを使用する場合には、拡張論理デバイスを構成する論理デバイス番号のリストを指定し、拡張論理デバイスを使用しない場合には未定義を指定する。例えば、図 3 では、論理デバイス番号 5 は論理デバイス番号 5、6、7 から構成され、各論理デバイスのサイズの合計である 12GB の論理デバイスとしてホスト 100 に提供される。

障害発生時データ退避情報 209 は、障害発生時において、各論理デバイスに関するキャッシュメモリ 112 上のダーティデータがディスク装置 114 に完全に退避されたかを管理するための情報である。状態としては、「未定義」「未完了」「完了」が存在する。「未定義」は、障害が発生していないことを示す。「未完了」は障害が発生しているが、キャッシュメモリ 112 上のダーティデータの退避が完了していないことを示す。「完了」は障害が発生しているが、キャッシュメモリ 112 上のダーティデータの退避が完了していることを示す。障害発生時データ退避情報 209 には、通常時は「未定義」が設定されているが、障害が検出された時点で、キャッシュメモリ 112 上にダーティデータを有する全ての論理デバイスに関して「未完了」が設定される。

【0030】

次に、LUPAS 管理テーブル 3 の一実施例について図 4 を用いて説明する。LUPAS 管理テーブル 3 は、ポート番号 301、ターゲット ID/LUN 302、接続ホスト番号 303、対応論理デバイス番号 304 を含む。ストレージ 110 内の各ポート 121 につき、有効な LUN 分の情報を保持する。ターゲット ID/LUN 302 は、ポート 121 に定義された LUN のアドレスを格納する。接続ホスト番号 303 は、当該ポート 121 の当該 LUN に対してアクセスを許可されているホスト 100 を示す情報である。一つの論理デバイスに対して複数ポート 121 の LUN が定義されている場合、それら全 LUN の接続ホスト番号の和集合が論理デバイス管理情報 2 の接続ホスト番号 205 に保持される。対応論理デバイス番号 304 には、当該 LUN を割り当てた論理デバイスの番号を格納する。例えば、図 4 では、ポート番号 1、2、…に対して、ターゲット ID に 0、LUN に 0、接続ホスト番号に 1、対応論理デバイス番号に 1 を指定している。

【0031】

次に、物理デバイス管理テーブル 4 の一実施例について図 5 を用いて説明する。物理デ

バイス管理テーブル4は、物理デバイス番号401、サイズ402、対応論理デバイス番号リスト403、デバイス状態404、RAID構成405、ディスク番号リスト406、ストライプサイズ407、ディスク内サイズ408、ディスク内開始オフセット409を含む。サイズ402には、物理デバイス番号401により特定される物理デバイスの容量が格納されている。対応論理デバイス番号リスト403には、当該物理デバイスが対応するストレージ110内の論理デバイス番号のリストが格納される。論理デバイスへ未割り当ての場合、当該エントリには無効値が設定される。デバイス状態404には、当該物理デバイスの状態を示す情報が設定される。状態としては、「オンライン」、「オフライン」、「未実装」、「障害オフライン」が存在する。「オンライン」は、当該物理デバイスが正常に稼動し、論理デバイスに割り当てられている状態であることを示す。「オフライン」は、当該物理デバイスは定義され、正常に稼動しているが、論理デバイスに未割り当てであることを示す。「未実装」は、当該物理デバイスがディスク装置114上に定義されていない状態にあることを示す。「障害オフライン」は、当該物理デバイスに障害が発生して論理デバイスに割り当てられないことを示す。なお、本実施形態では、簡単のため、物理デバイスは製品の工場出荷時にあらかじめディスク装置114上に作成されているものとする。このため、利用可能な物理デバイスについてはデバイス状態404の初期値は「オフライン」状態、その他は「未実装」状態となる。RAID構成405には、当該物理デバイスが割り当てられたディスク装置114のRAIDレベル、データディスクとパリティディスク数などRAID構成に関連する情報が保持される。同じように、ストライプサイズ407には、RAIDにおけるデータ分割単位(ストライプ)長が保持される。ディスク番号リスト406には、当該物理デバイスが割り当てられたRAIDを構成する複数のディスク装置114の番号が保持される。この番号はストレージ110内でディスク装置114を識別するために付与した一意な値である。ディスク内サイズ408とディスク内開始オフセット409には、当該物理デバイスデータが各ディスク装置114内のどの領域に割り当てられているかを示す情報である。本実施例では簡単のため、全物理デバイスについてRAIDを構成する各ディスク装置114内のオフセットとサイズを統一している。

【0032】

次に、スロット管理テーブル5の一実施例について図6を用いて説明する。ホスト100は、ストレージ110内の論理デバイスのデータを指定するために論理アドレスを使用する。論理アドレスは、例えば、論理デバイス番号と論理デバイス内の位置情報とからなる。ストレージ110内では、連続した論理アドレス空間を一定量に分割して管理しており、分割した論理アドレスをスロットと呼ぶ。論理アドレス空間の分割のサイズはスロットサイズとよぶ。例えば、スロット番号は、論理アドレス空間をスロットサイズで割った商に1を加算したものとする。スロット管理テーブル5はスロット番号501、セグメント番号リスト502、スロット属性503、論理デバイス番号504、ホスト番号505、ロック情報506を含む。セグメント番号リスト502には、当該スロットが含むセグメント(分割したキャッシュメモリ)のセグメント番号がリスト形式で保持される。例えば、図6では、ひとつのスロットが4つのセグメントから構成される例を示している。当該スロットの対応する位置のセグメントが存在しない場合、無効値として0を指定する。スロット属性503には、当該スロットの属性が保持される。属性としては「クリーン」「ダーティ」「フリー」が存在する。「クリーン」は当該スロットが保持するキャッシュメモリ112上のデータがディスク装置114上のデータと一致することを意味する。「ダーティ」は当該スロットが保持するキャッシュメモリ112上のデータがディスク装置114に未反映であることを意味する。「フリー」は当該スロットが使用されていないことを意味する。論理デバイス番号504には、当該スロットに対応する論理デバイス番号が保持される。ホスト番号505には、当該スロットに対応する入出力要求を行ったホスト100のホスト番号が保持される。ロック情報506には当該スロットのチャネルアダプタ120およびディスクアダプタ130間での排他のためのロック情報が保持される。状態としては「オン」「オフ」が存在し、「オン」は当該スロットがロックされていること

を意味し、「オフ」は当該スロットがロックされていないことを意味する。

【0033】

次に、セグメント管理テーブル6の一実施例について図7を用いて説明する。セグメント管理テーブル6は、セグメント番号601、ブロック情報602を含む。ブロック情報602は、ホスト100からのアクセスの単位（以下、ブロックよぶ）で、セグメントの当該位置に保持しているデータが有効か無効を示す。図7に示すセグメント管理テーブルの例では、セグメントサイズが2048バイト、ブロックサイズが512バイトの場合であり、セグメント管理テーブル6は4つのブロック情報を持ち、セグメント番号1のブロック位置1とブロック位置3が有効、すなわち、セグメントの先頭から512バイト分と1024バイトから512バイト分に有効なデータがあることを示す。

【0034】

次に、チャネルジョブ管理テーブル7の一実施例について図8を用いて説明する。チャネルジョブとは、チャネルアダプタ120上で動作するジョブのことを指す。チャネルジョブ管理テーブル7は、ジョブ番号701、処理702、論理デバイス番号703、転送開始位置704、転送長705、ホスト番号706を含む。処理702には当該ジョブの処理内容が含まれる。処理内容には「リード」「ライト」が存在する。論理デバイス番号703は当該ジョブの処理対象となる論理デバイス番号を保持する。転送開始位置704は当該ジョブの処理対象となる論理デバイス上のアドレスを保持する。転送長705は当該ジョブの処理対象となるアクセス長を保持する。ホスト番号706は当該ジョブの処理に対応するホスト100のホスト番号が保持される。

【0035】

次に、ディスクジョブ管理テーブル8の一実施例について図9を用いて説明する。ディスクジョブとは、ディスクアダプタ130上で動作するジョブのことを指す。ディスクジョブ管理テーブル8は、ジョブ番号801、処理802、論理デバイス番号803、転送開始位置804、転送長805、ホスト番号806、チャネルアダプタ番号807を含む。内容に関しては図8で述べたものと同様なものは、説明を省略する。チャネルアダプタ番号807には当該ジョブの要求元に対応するチャネルアダプタのチャネルアダプタ番号が保持される。

【0036】

次に、ホスト管理テーブル9の一実施例について図10を用いて説明する。ホスト管理テーブル9は、ホスト番号901、ホスト名/WWN902、ホスト優先度903を含む。ホスト名/WWN902は、ホストを一意に識別するための情報である。ホスト優先度903は、ホスト100での入出力処理の重要度を反映するものであり、チャネルアダプタおよびディスクアダプタ上のジョブをスケジューリングする際の優先度として用いられる。優先度が高いジョブとは、例えば、無停止運転のオンライントランザクション処理などの高速な応答が要求される処理を行うジョブであり、また、優先度の低いジョブとは、例えば、夜間バッチ処理のような高速な応答が要求されない処理を行うジョブのことである。

【0037】

次に、アクセスパターン管理テーブル10の一実施例について図11を用いて説明する。アクセスパターン管理テーブル10は、論理デバイス番号1001、リード回数1002、ライト回数1003、リードヒット回数1004、ライトヒット回数1005、シーケンシャルリード回数1006、シーケンシャルライト回数1007を含む。リード回数1002は当該論理デバイスに対するリード回数を示す。ライト回数1003は当該論理デバイスに対するライト回数を示す。リードヒット回数1004は当該論理デバイスに対するリードヒット回数を示す。ライトヒット回数1005は当該論理デバイスに対するライトヒット回数を示す。シーケンシャルリード回数1006は当該論理デバイスに対するシーケンシャルリード回数を示す。シーケンシャルライト回数1007は当該論理デバイスに対するシーケンシャルライト回数を示す。ダーティ量管理情報1008は、各論理デバイスに関してキャッシュメモリ112上のダーティデータの量に関する情報を保持する。

ものであり、例えば、ダーティデータ量の過去 24 時間の平均値、過去 1 時間の平均値、現在値を保持する。

【0038】

次に、業務管理テーブル 11 の一実施例について図 12 を用いて説明する。本実施の形態において、業務とは、ホストと論理デバイスの組により定義される。一般的に言う業務は前記定義と規模が異なる場合があるが、本実施の形態では一例として前記定義のようにした。業務管理テーブル 11 は、業務番号 1101、論理デバイス番号 1102、ホスト番号 1103、業務優先度 1104 を含む。論理デバイス番号 1102、ホスト番号 1103 は業務に対応する論理デバイスとホストの組を示す。業務優先度 1104 は、ストレージ 110 での障害発生時におけるデータの優先度を示す情報であり、例えば、最高優先度の値 1 から最低優先度の値 5 までの数値を用いる。例えば、図 12 では、業務番号 1 に関して、論理デバイス番号に 0、ホスト番号に 0、業務優先度に 1 を指定している。本優先度に基づきチャネルアダプタおよびディスクアダプタ上のジョブのスケジューリングが行われる。ジョブの優先度の高さの意味の例はホスト管理テーブル 9 で説明したものと同様である。

【0039】

次に、スケジューリング管理テーブル 12 の一実施例について図 13 を用いて説明する。スケジューリング管理テーブル 12 は障害発生時ジョブスケジューリング時のスケジューリングパラメータを管理するものであり、論理デバイス優先度 1201、ホスト優先度 1202、業務優先度 1203 の 3 つの優先度に対して、スケジューリングの比率を指定する。本比率は後述するチャネルアダプタにおけるジョブスケジューリング処理（図 20 に示す処理）、チャネルアダプタにおけるジョブスケジューリング処理（図 23 に示す処理）に用いられる。例えば、図 13 では、論理デバイス優先度、ホスト優先度、業務優先度にそれぞれ、0.5、0.3、0.2 の比率を設定しており、ジョブスケジューリング処理において、前記論理デバイス優先度は 10 回に 5 回の割合で使用される。

【0040】

次に、ストレージ管理者が管理端末 ST（140）を用いて論理デバイス定義処理 13 及び LU パス定義処理 14 並びに論理デバイス優先度定義処理 15、ホスト優先度定義処理 16 及び業務優先度定義処理 17 を行うことについて図 14 及び図 15 並びに図 16、図 17 及び図 18 を用いて説明する。まず、ストレージ管理者が管理端末 ST（140）を用いて論理デバイス定義処理 13 及び LU パス定義処理 14 を行うことによりホスト 100 はストレージ S（110）内の論理デバイスにアクセスできるようになる。その後、ストレージ管理者が管理端末 ST（140）を用いて論理デバイス優先度定義処理 15、ホスト優先度定義処理 16 及び業務優先度定義処理 17 を行うことにより、ストレージ S（110）は設定された各優先度（論理デバイス優先度、ホスト優先度及び業務優先度）に基づき適切に障害発生時の動作をすることが可能となる。以下、図中で、ST は管理端末 140、S はストレージ 110 を表す。

【0041】

まず、管理端末 140 を用いた論理デバイス定義処理の一実施例について図 14 を用いて説明する。論理デバイス定義処理 13 は、ストレージ管理者による管理端末 140 からの指示を受け付けて、ストレージ 110 が保持するディスク装置 114 に対して論理デバイスを定義する処理である。まず、管理端末 140 の CPU 145 は、ストレージ管理者より入力装置 141 やディスプレイ 142 等を用いて入力された論理デバイスの定義（論理デバイス番号及び物理デバイス番号等）を受け付けてメモリ 143 等に記憶させる（ステップ 1301）。その後、管理端末 140 の CPU 145 が、ストレージ管理者より、論理デバイスの定義として配置先の物理デバイス番号の指定の入力があったかを判定し（ステップ 1302）、特別な理由があつて配置先の物理デバイス番号の指定があつた場合には、管理端末 140 の CPU 145 はそのまま I/F 146、119 を介してストレージ 110 に対して論理デバイスの定義を指示する（ステップ 1305）。管理端末 140 より指示を受け付けたストレージ 110 は、図 3 に示す論理デバイス管理テーブル 2 に論

理デバイス番号、サイズ、デバイス状態、物理デバイス番号、開始アドレス、論理デバイス優先度を設定する。論理デバイス優先度 207 にはデフォルト値、例えば 3 を設定する。その後、管理端末 140 に完了報告を送信する（ステップ 1306）。デバイス状態としては初期値である「オフライン」が設定される。最後に、管理端末 140 は、ストレージ 110 より完了報告を受け付ける（ステップ 1307）。

【0042】

配置先の物理デバイス番号の指定がなかった場合には、管理端末 140 の CPU 145 は、共有メモリ 113 に記憶された図 3 に示す論理デバイス管理テーブル 2 及び図 5 に示す物理デバイス管理テーブル 4 の複製を管理端末 140 のメモリ 143 に格納し、該格納された情報を参照することにより（ステップ 1303）、配置先の物理デバイスを決定し、ストレージ 110 に転送されて共有メモリ 113 内の論理デバイス管理テーブル 2 及び物理デバイス管理テーブル 4 に設定されることになる。具体的には、優先度が高い論理デバイスのキャッシュメモリ 112 上のデータを早急にディスク装置 114 に退避できるように、配置先の物理デバイスを決定し、ストレージ 110 に転送されて共有メモリ 113 内の論理デバイス管理テーブル 2 及び物理デバイス管理テーブル 4 に設定されることになる。決定方法としては、例えば、同一優先度の論理デバイスを異なる複数の物理デバイス（1 又は複数のディスク装置 114 から構成される）に配置する。これにより、データ退避時に物理デバイスへのアクセス競合が起きず、高速なデータ退避が期待できる。また、別の決定方法としては、前述の拡張論理デバイスを用いることにより、一つの論理デバイスを複数の論理デバイスから構成する。このように各論理デバイスを配置する物理デバイスを分散させることにより、データ退避時に物理デバイスへのアクセス競合が起きず、高速なデータ退避が期待できる。その後、ストレージ管理者に対して管理端末 140 のディスプレイ 142 を介して論理デバイスの配置先として最適な（推奨する）物理デバイスを提示する（ステップ 1304）。その後の処理は、ステップ 1305 以降と同様である。

【0043】

即ち、論理デバイス定義処理 13 は、管理端末 140 はホストに提供する論理デバイスに対する論理デバイスの優先度（第一の優先度）を受け付け、該論理デバイスの優先度に基づき 1 又は複数のディスク装置から構成される複数の物理デバイスに、複数の論理デバイスを分散して配置するように（第一の優先度の高い論理デバイスは第一の優先度の低い論理デバイスよりも多くの物理デバイスを割り当てるように）、ストレージ 110 に転送して共有メモリ 113 内の論理デバイス管理テーブル 2 及び物理デバイス管理テーブル 4 に設定する処理である。従って、論理デバイス定義処理 13 を予め行うことによって、ストレージ 110 の制御装置 130 は、障害発生時にディスク装置に未反映なキャッシュメモリ 112 上のホスト 100 から書き込まれたデータであるダーティデータのうちの重要なデータを速やかにディスク装置に反映することが可能となる。

【0044】

次に、管理端末 140 を用いた LU パス定義処理の一実施例について図 15 を用いて説明する。LU パス定義処理 14 は、ストレージ管理者による管理端末 140 からの指示を受け付けて、ストレージ 110 が提供する論理デバイスをホスト 100 からアクセス可能な状態に設定する処理である。まず、管理端末 140 の CPU 145 は、入力装置 141 等を用いて LU パス定義指示を受け付け、該受け付けた同指示をストレージ 110 に転送する（ステップ 1401）。同指示には、LU を定義するポート番号、LUN、接続ホスト番号、対象となる論理デバイス番号が含まれる。ストレージ 110 は同指示に従い、図 4 に示す LU パス管理テーブル 3 の各エントリに値を設定後、管理端末 140 に完了報告を送信する（ステップ 1402）。最後に、管理端末 140 は、ストレージ 110 より完了報告を受け付ける（ステップ 1403）。

【0045】

次に、管理端末 140 を用いた論理デバイス優先度定義処理の一実施例について図 16 を用いて説明する。論理デバイス優先度定義処理 15 は、障害発生時にキャッシュメモリ 112 上のダーティデータをディスク装置 114 に反映する順序を定義するための処理であ

る。まず、管理端末140のCPU145は、ストレージ管理者から入力装置141を用いて論理デバイスの優先度を受け付ける。入力パラメータには、論理デバイス番号と論理デバイス優先度を含む。管理端末140は、その後、入力パラメータをストレージ110に送信する(ステップ1501)。管理端末140より入力パラメータを受け取ったストレージ110は、入力パラメータに基づき、図3に示す論理デバイス管理テーブル2の論理デバイス番号と論理デバイス優先度を設定し、管理端末140に登録が完了したことを通知する(ステップ1502)。ストレージ110から登録完了報告を受けた管理端末140はストレージ管理者に完了報告を行う(ステップ1503)。従って、ストレージ110の制御装置130は、障害発生時に、上記論理デバイス優先度定義処理15によって設定された論理デバイスの優先度に基づいた順序で、キャッシュメモリ112上のダーティデータをディスク装置114に反映する処理を行うことが可能となる。

【0046】

次に、管理端末140を用いたホスト優先度定義処理の一実施例について図17を用いて説明する。ホスト優先度定義処理16は、ホスト100からの入出力要求の処理優先度を定義する処理である。まず、管理端末140のCPU145は、ストレージ管理者から入力装置141により、ホスト100の優先度を受け付ける。入力パラメータには、ホスト番号と優先度を含む。管理端末140は、その後、入力パラメータをストレージ110に送信する(ステップ1601)。管理端末140より入力パラメータを受け取ったストレージ110は、入力パラメータに基づき、図10に示すホスト管理テーブル9のホスト優先度903を設定し、管理端末140に登録が完了したことを通知する(ステップ1602)。ストレージ110から登録完了報告を受けた管理端末140はストレージ管理者に完了報告を行う(ステップ1603)。

【0047】

次に、管理端末140を用いた業務優先度定義処理の一実施例について図18を用いて説明する。まず、管理端末140のCPU145は、ストレージ管理者から入力装置141により、障害発生時に重要業務への性能低下を回避するための、業務の優先度(第二の優先度)を受け付ける。入力パラメータには、業務番号、論理デバイス番号とホスト番号の組、業務優先度を含む。管理端末140は、その後、入力パラメータをストレージ110に送信する(ステップ1701)。管理端末140より入力パラメータを受け取ったストレージ110は、入力パラメータに基づき、図12に示す業務管理テーブル11の論理デバイス番号1102、ホスト番号1103、業務優先度1104を設定し、管理端末140に登録が完了したことを通知する(ステップ1702)。ストレージ110から登録完了報告を受けた管理端末140はストレージ管理者に完了報告を行う(ステップ1703)。このように、ストレージ110は、管理端末140によって業務の優先度(第二の優先度)が設定されるので、障害発生時に設定された業務の優先度を用いてストレージ内のジョブをスケジューリングする処理を行うことが可能となり、その結果重要業務への性能低下を回避することが可能となる。なお、業務優先度は、ホスト優先度であってもよい。

【0048】

次に、チャネルアダプタ120におけるチャネルアダプタポート処理の一実施例について図19を用いて説明する。チャネルアダプタポート処理18では、チャネルアダプタ120は、ポート121を介してホスト100からのコマンドを受け付け、図8に示すチャネルジョブ管理テーブル7にジョブを登録する。本処理においては、ジョブを図示しないFIFOキュー、論理デバイスプライオリティーキュー、ホストプライオリティキュー、業務プライオリティーキューにエンキューする(ステップ1801)。プライオリティーキューは平衡木やB木などの公知のデータ構造により実現される。論理デバイスプライオリティーキュー、ホストプライオリティキュー、業務プライオリティーキューにエンキューする際の優先度の値としては、それぞれ、図3に示す論理デバイス管理テーブル2における、ホスト100からのリード・ライトコマンドに対応する論理デバイスの論理デバイス優先度207、図10に示すホスト管理テーブル9におけるホスト優先度903、図12に示す業務管理テーブル11における業務優先度1104を用いる。

【0049】

次に、本発明に係る制御装置であるチャンネルアダプタ120におけるジョブスケジューリング処理の一実施例について図20を用いて説明する。チャンネルアダプタ120におけるジョブスケジューリング処理19では、チャンネルアダプタ120は、まず、障害が発生している状態であるかを判定し（ステップ1901）、障害が発生していなければ、図8に示すチャンネルジョブ管理テーブル7のジョブが接続されている図示しないFIFOキューの端に位置するジョブをデキューし、当該ジョブを実行する（ステップ1902）。この際、当該ジョブは、図示しない論理デバイスプライオリティーキュー、ホストプライオリティーキュー、業務プライオリティーキューからもデキューされる。チャンネルアダプタ120は、障害が発生している場合には、図8に示すチャンネルジョブ管理テーブル7のジョブが接続されている図示しない論理デバイスプライオリティーキュー、ホストプライオリティーキュー、業務プライオリティーキューのいずれかのプライオリティーキューの最も優先度が高いジョブをデキューし、当該ジョブを実行する（ステップ1903）。プライオリティーキューの選択は、図13に示すスケジューリング管理テーブル12における論理デバイス優先度1201、ホスト優先度1202、業務優先度1203の比率（管理端末140が設定する。）を用いる。また、この際、当該ジョブは、図示しないFIFOキューからもデキューされる。ただし、プライオリティーキューのみによりジョブのスケジューリングを行うと、優先度の高いジョブのみが実行され、優先度の低いジョブが長期間実行されない可能性があるので、一定の割合で、プライオリティーキューのかわりにFIFOキューを用いる。以上説明したチャンネルアダプタ120におけるジョブスケジューリング処理19の中で実行されるジョブは、図21に示すリードジョブ処理20もしくは図22に示すライトジョブ処理21を実行することになる。

【0050】

次に、チャンネルアダプタ120におけるリードジョブ処理の一実施例について図21を用いて説明する。チャンネルアダプタ120におけるリードジョブ処理20では、キャッシュメモリ112上にホスト100からのリード要求対象データがあれば当該データをホスト100に転送し、なければディスクアダプタ130にリード要求対象データのステージングを要求する。まず、チャンネルアダプタ120は、リード要求対象論理デバイスのデバイス状態をチェックし（ステップ2002）、「オンライン」以外であれば（エラー）、ホスト100にエラーを送信し、処理を終了する。チャンネルアダプタ120は、エラーでなければ、ジョブスケジューリング処理19からの要求を解析し、リード要求対象データに対応するスロット番号、セグメント位置、ブロック位置を算出する（ステップ2002）。その後、スロット番号に対応するスロットの参照および更新を行うが、他のチャンネルアダプタ120およびディスクアダプタ130が当該スロットに同時にアクセスしないように当該スロットのロックを取得する（ステップ2003）。具体的には、図6に示すスロット管理テーブル5のロック情報に「オン」を設定する。以後、ロックの取得の処理はここで説明した処理と同様であるので、説明を省略する。次に、リード要求対象データのヒットミス判定を行う（ステップ2004）。具体的には、スロット管理テーブル5より、対象となるスロット番号に対応するセグメント番号リストを参照し、対象となるセグメント位置のセグメント番号を得る。次に、セグメント管理テーブル6より、対象となるセグメント番号に対応するブロック情報を参照し、対象となるブロック位置のデータが有効か無効かを判定する。

【0051】

チャンネルアダプタ120は、ステップ2004において当該データが有効な場合（ヒット）には、まず、アクセス情報を更新する（ステップ2005）。具体的には、リード要求対象データに対応する、アクセスパターン管理テーブル10における論理デバイスのリード回数とリードヒット回数を1だけインクリメントする。また、図示しない共有メモリ113上のシーケンシャル学習情報（連続する領域に対するリードアクセス情報）より、当該リード要求がシーケンシャルリードであるかを判定する。シーケンシャルリードとは、ホスト100から論理デバイスに対して論理デバイスのアドレス空間における連続した

一連のリードアクセスである。シーケンシャルリードを検出した際、ホスト100からの要求とは非同期に当該論理デバイスに対する最後のリードアクセスのデータ以降のデータを先読みする。これにより、同期リード処理において、リード要求がヒットする確率が高くなり、高速なリードアクセスが期待できる。シーケンシャルリードであれば、リード要求対象データに対応する、アクセスパターン管理テーブル10における論理デバイスのシーケンシャルリード回数を1だけインクリメントする。その後、当該スロットをクリーンキューのMRUの端に遷移させる（ステップ2006）。次に、シーケンシャル学習情報に基づき、当該リード要求がシーケンシャルリードであれば、先読みのために、ディスクジョブ管理テーブル8に1または複数のジョブを登録する（ステップ2007）。具体的には、図19において説明した手順と同様にFIFOキュー、論理デバイスプライオリティーキュー、ホストプライオリティーキュー、業務プライオリティーキューにジョブをエンキューする。最後に、ホスト100に対象となるデータを転送し（ステップ2008）、スロットのロックを開放する（ステップ2009）。具体的には、スロット管理テーブル5のロック情報に「オフ」を設定する。以後、ロックの解放の処理はここで説明した処理と同様であるので、説明を省略する。以上がヒットの場合の処理である。

【0052】

チャネルアダプタ120は、ステップ2004において当該データが無効な場合（ミス）には、まず、アクセス情報を更新する（ステップ2010）。具体的には、リード要求対象データに対応する、アクセスパターン管理テーブル10における論理デバイスのリード回数を1だけインクリメントする。アクセスパターン管理テーブル10のシーケンシャルリード回数の更新はヒットの場合と同様に行う。次に、必要数のキャッシュセグメントの新規確保を行う（ステップ2011）。キャッシュセグメントの新規確保は、図示しない使用されていないキャッシュセグメントを管理するキュー（フリーキュー）に接続されているキャッシュセグメント、もしくは、LRUアルゴリズムのような公知な技術によりキュー管理されている、スロット属性が「クリーン」であるスロットが保持するキャッシュセグメントを用いて行う。次に、ディスクジョブ管理テーブル8にジョブを登録する（ステップ2012）。具体的には、図19において説明した手順と同様にFIFOキューとプライオリティーキューにジョブをエンキューする。次に、シーケンシャル学習情報に基づき、先読み処理を行うが、具体的な手順はヒット時で説明した手順と同様である（ステップ2013）。その後、スロットのロックを解放し（ステップ2014）、ディスクアダプタ130でステージングが完了するまで待つ（ステップ2015）。ディスクアダプタ130よりステージングの完了報告を受信（ステップ2016）したのち、再びステップ2003の処理から再開する。この後のヒットミス判定処理においては、すでにステージングが完了しているため、以後の処理はヒットの場合と同様になる。

【0053】

次に、チャネルアダプタ120におけるライトジョブ処理の一実施例について図22を用いて説明する。チャネルアダプタ120におけるライトジョブ処理21では、ホスト100からライト要求対象データを受信しキャッシュメモリ112に格納した後、必要であればディスクアダプタ130に対象データのデステージングを要求し、ホスト100に完了報告を行う。まず、チャネルアダプタ120は、ライト要求対象論理デバイスのデバイス状態をチェックし（ステップ2101）、「オンライン」以外であれば（エラー）、ホスト100にエラーを送信し、処理を終了する。チャネルアダプタ120は、エラーでなければ、ジョブスケジューリング処理19からの要求を解析し、ライト要求対象データに対応するスロット番号、セグメント位置、ブロック位置を算出する（ステップ2102）。その後、図21で説明した理由と同様の理由により、スロット番号に対応するスロットのロックを取得する（ステップ2103）。次に、図21で説明した手順と同様にライト要求対象データのヒットミス判定を行う（ステップ2104）。

【0054】

チャネルアダプタ120は、ヒットの場合には、まず、アクセス情報を更新する（ステップ2105）。具体的には、ライト要求対象データに対応する、アクセスパターン管理

テーブル10における論理デバイスのライト回数とライトヒット回数を1だけインクリメントする。また、図示しない共有メモリ113上のシーケンシャル学習情報より、当該ライト要求がシーケンシャルライトであるかを判定する。シーケンシャルライトとは、ホスト100から論理デバイスに対して論理デバイスのアドレス空間における連続した一連のライトアクセスである。シーケンシャルライトであれば、ライト要求対象データに対応する、アクセスパターン管理テーブル10における論理デバイスのシーケンシャルライト回数を1だけインクリメントする。次に、ホスト100に転送準備完了のメッセージを送信する（ステップ2106）。その後、ホスト100からライト要求対象データを受信し、キャッシュメモリ112に格納して、ダーティキューのMRUの端にキュー遷移を行う（ステップ2107）。

【0055】

チャンネルアダプタ120は、ミスの場合には、まず、アクセス情報を更新する（ステップ2108）。具体的には、ライト要求対象データに対応する、アクセスパターン管理テーブル10における論理デバイスのライト回数を1だけインクリメントする。アクセスパターン管理テーブル10のシーケンシャルライト回数の更新はヒットの場合と同様に行う。次に、必要数のキャッシュセグメントを新規確保（ステップ2109）した後に、ホスト100に転送準備完了のメッセージを送信する（ステップ2110）。その後、ホスト100からライト要求対象データを受信し、キャッシュメモリ112に格納して、ダーティキューのMRUの端にエンキューする（ステップ2111）。

【0056】

以後の処理は同期ライトが必要かどうかで処理が異なる（ステップ2112）。同期ライトが必要な場合とは、後述する共有メモリ113上の障害発生中フラグが「ON」のときであり、不要な場合とは、同フラグが「OFF」のときである。同期ライトが不要な場合は、ホスト100にライト完了報告を送信する（ステップ2117）。同期ライトが必要な場合は、ディスクジョブ管理テーブル8にジョブを登録する（ステップ2113）。具体的には、図19において説明した手順と同様にFIFOキューとプライオリティーキューにジョブをエンキューする。その後、スロットのロックを解放し（ステップ2114）、ディスクアダプタでデステージングが完了するまで待つ（ステップ2115）。ディスクアダプタ130よりデステージングの完了報告を受信（ステップ2116）した後、ホスト100に完了報告を送信する（ステップ2117）。同期ライト処理では、データがディスク装置114に確実に書き込まれることが保障される。

【0057】

以上がチャンネルアダプタ120によるジョブスケジューリング処理19である。

【0058】

次に、本発明に係る制御装置であるディスクアダプタ130におけるジョブスケジューリング処理の一実施例について図23を用いて説明する。ディスクアダプタ130におけるジョブスケジューリング処理22は、図20で説明した手順と異なりチャンネルジョブ管理テーブル7ではなくディスクジョブ管理テーブル8を用いる点と以下で述べる障害発生時のダーティデータに関する処理が追加されている点を除いては、図20の説明と同様の手順で処理を行う。即ち、ディスクアダプタ130は、まず、障害が発生している状態であるかを判定し（ステップ2201）、障害が発生していなければ、図9に示すディスクジョブ管理テーブル8のジョブが接続されている図示しないFIFOキューの端に位置するジョブをデキューし、当該ジョブを実行する（ステップ2202）。ディスクアダプタ130は、障害が発生している場合には、図9に示すディスクジョブ管理テーブル8のジョブが接続されている図示しない論理デバイスプライオリティーキュー、ホストプライオリティーキュー、業務プライオリティーキューのいずれかのプライオリティーキューの最も優先度が高いジョブをデキューし、当該ジョブを実行する（ステップ2203）。そして、ディスクアダプタ130は、ステップ2204では、障害発生時に、論理デバイス管理テーブル2の障害発生時データ退避情報209をチェックし、キャッシュメモリ112上のダーティデータがディスク装置114に完全に退避されていないことを示す「未完了」

が設定されている論理デバイスに関して、アクセスパターン管理テーブル10のデータ量管理情報1008のデータ量の現在値を参照し、値がゼロであれば、障害発生時データ退避情報209に「完了」を設定する。以上説明したディスクアダプタ130におけるジョブスケジューリング処理22の中で実行されるジョブは、図25に示すリードジョブ処理24もしくは図26に示すライトジョブ処理25を実行することになる。

【0059】

次に、ディスクアダプタ130における非同期ライトジョブ登録処理の一実施例について図24を用いて説明する。ディスクアダプタ130における非同期ライトジョブ登録処理23では、キャッシュメモリ112に格納されたライトデータを物理デバイスに書き込む。まず、データキューのLRUの端より対象となるスロットを選択し、データキューからデキューする（ステップ2301）。次に、ディスクジョブ管理テーブル8にジョブを登録する（ステップ2302）。具体的には、図19において説明した手順と同様にFIFOキューとプライオリティーキューにジョブをエンキューする。

【0060】

次に、ディスクアダプタ130におけるリードジョブ処理の一実施例について図25を用いて説明する。ディスクアダプタ130におけるリードジョブ処理24では、まず、ジョブスケジューリング処理22からの要求を解析し、リード要求対象データに対応するスロット番号、セグメント位置、ブロック位置を算出する（ステップ2401）。その後、スロット番号に対応するスロットの参照および更新を行うが、他のディスクアダプタ130およびチャネルアダプタ120が当該スロットに同時にアクセスしないように当該スロットのロックを取得する（ステップ2402）。次に、物理デバイスからリード要求対象データを読み出し、キャッシュメモリ112に格納し、クリーンキューのMRUの端にエンキュー（ステップ2403）した後、スロットのロックを解放し（ステップ2404）、ディスクジョブ管理テーブル8のチャネルジョブ番号807で指定されたチャネルジョブに対応するチャネルアダプタ120に対してステージング完了報告を送信する（ステップ2405）。

【0061】

次に、ディスクアダプタ130におけるライトジョブ処理の一実施例について図26を用いて説明する。ディスクアダプタ130におけるライトジョブ処理25では、まず、ジョブスケジューリング処理22からの要求を解析し、ライト要求対象データに対応するスロット番号、セグメント位置、ブロック位置を算出する（ステップ2501）。次に、当該スロットのロックを取得し（ステップ2502）、当該スロットが保持するデータデータを物理デバイスに書き込む（ステップ2503）。その後、当該スロットのスロット属性を「クリーン」に更新し、クリーンキューにエンキューする（ステップ2504）。最後に、当該スロットのロックを解放し（ステップ2505）、当該ライトジョブが同期ライトであればディスクジョブ管理テーブル8のチャネルジョブ番号807で指定されたチャネルジョブに対応するチャネルアダプタ120に対してデステージング完了報告を送信する（ステップ2506）。

【0062】

以上がディスクアダプタ130によるジョブスケジューリング処理22である。

【0063】

次に、本発明に係る障害処理の一実施例について図27を用いて説明する。障害処理26では、障害発生検出時に、制御装置120、130が各論理デバイスに関するキャッシュメモリ112上のデータデータを物理デバイスに反映するための処理である。まず、ディスクアダプタ130は障害処理を終了するかを判定する（ステップ2601）。具体的には、ストレージ110の動作を停止する場合などに処理を終了する。処理を終了しない場合、ディスクアダプタ130はストレージ110内の各部位に障害が発生しているかをチェックする（ステップ2602）。次に、ディスクアダプタ130は、障害が発生しているかを判定し（ステップ2603）、障害発生を検出した場合は、図3で説明したように障害発生時データ退避情報209を設定し、図示しない共有メモリ113上の障害発

生中フラグを「ON」に設定する(ステップ2604)。ここで、障害発生中フラグは通常時、「OFF」に設定されている。その後、チャンネルアダプタ120は、各論理デバイスのダーティデータに関するライトジョブのエンキュー処理を行う(ステップ2605)。具体的には、チャンネルアダプタ120は、各論理デバイスに関してアクセスパターン管理テーブル10のダーティ量管理情報1008の現在値を参照し、値がゼロでない論理デバイスについて、当該論理デバイスの論理アドレス空間を走査し、スロット属性が「ダーティ」であるスロットを検索する。そして、当該スロットに対応するライトジョブをディスクジョブ管理テーブル8に登録する。具体的には、図19において説明した手順と同様にFIFOキューとプライオリティーキューにジョブをエンキューする。ディスクアダプタ130が、ステップ2603において、障害発生を検出しなかった場合は、ステップ2601に戻る。チャンネルアダプタ120は、ステップ2605の後、対象ライトジョブの完了を待つ。ディスクアダプタ130がこの過程で予備電源不足もしくは対象キャッシュデータに関してキャッシュメモリの二重障害を検出(ステップ2606)した場合には、対象論理デバイスのデータは消失したことになるので、チャンネルアダプタ120は対象論理デバイスのデバイス状態を「障害オフライン」に設定する(ステップ2607)。ステップ2606においてすべてのライトが完了した場合にはステップ2601に戻る。

【0064】

本発明に係る別の実施の形態としては、ストレージ110に障害発生後、図27に示す障害処理26によるライトデータの反映が終了しなかった場合もしくはキャッシュに二重障害が発生した場合に、キャッシュメモリ112上のダーティデータが消失した際、当該データのスロット位置を管理端末140のディスプレイ142に表示するものである。本実施形態においては、論理デバイス管理テーブル2を参照し、障害時データ退避情報が「未完了」である論理デバイスに対応するダーティデータのスロット位置をスロット管理テーブル5を参照することにより算出し、該当するスロット番号をすべて論理デバイス番号と共にディスプレイ142に表示する。これにより、ストレージ管理者は消失したデータ領域にもをリカバリすることで、リカバリ時間の短縮が可能となる。

【0065】

以上説明したように本実施の形態によれば、ストレージに障害が発生した際に、ストレージ内のキャッシュメモリ上のダーティデータを速やかにディスク装置に反映することができ、優先度の高い重要なデータの消失を回避することができる。また、ストレージに障害が発生した際に、優先度の高い重要な業務に関する性能低下を極力回避することができる。

【図面の簡単な説明】

【0066】

【図1】本発明に係るストレージを含む計算機システムの一実施の形態を示す構成図である。

【図2】本発明に係る共有メモリに記憶される各種管理テーブルを示す図である。

【図3】本発明に係る論理デバイス管理テーブルの一実施例を示す構成例示図である。

。

【図4】本発明に係るLUパス管理テーブルの一実施例を示す構成例示図である。

【図5】本発明に係る物理デバイス管理テーブルの一実施例を示す構成例示図である。

。

【図6】本発明に係るスロット管理テーブルの一実施例を示す構成例示図である。

【図7】本発明に係るセグメント管理テーブルの一実施例を示す構成例示図である。

【図8】本発明に係るチャンネルジョブ管理テーブルの一実施例を示す構成例示図である。

【図9】本発明に係るディスクジョブ管理テーブルの一実施例を示す構成例示図である。

【図10】本発明に係るホスト管理テーブルの一実施例を示す構成例示図である。

【図11】本発明に係るアクセスパターン管理テーブルの一実施例を示す構成例示図

である。

【図 12】本発明に係る業務管理テーブルの一実施例を示す構成例示図である。

【図 13】本発明に係るスケジューリング管理テーブルの一実施例を示す構成例示図である。

【図 14】本発明に係る論理デバイス定義処理の一実施例を示す処理フロー図である。

【図 15】本発明に係る L U パス定義処理の一実施例を示す処理フロー図である。

【図 16】本発明に係る論理デバイス優先度定義処理の一実施例を示す処理フロー図である。

【図 17】本発明に係るホスト優先度定義処理の一実施例を示す処理フロー図である。

【図 18】本発明に係る業務優先度定義処理の一実施例を示す処理フロー図である。

【図 19】本発明に係るチャンネルアダプタでのチャンネルアダプタポート処理の一実施例を示す処理フロー図である。

【図 20】本発明に係るチャンネルアダプタでのジョブスケジューリング処理の一実施例を示す処理フロー図である。

【図 21】本発明に係るチャンネルアダプタでのリードジョブ処理の一実施例を示す処理フロー図である。

【図 22】本発明に係るチャンネルアダプタでのライトジョブ処理の一実施例を示す処理フロー図である。

【図 23】本発明に係るディスクアダプタでのジョブスケジューリング処理の一実施例を示す処理フロー図である。

【図 24】本発明に係る非同期ライトジョブ登録処理の一実施例を示す処理フロー図である。

【図 25】本発明に係るディスクアダプタでのリードジョブ処理の一実施例を示す処理フロー図である。

【図 26】本発明に係るディスクアダプタでのライトジョブ処理の一実施例を示す処理フロー図である。

【図 27】本発明に係る障害処理の一実施例を示す処理フロー図である。

【符号の説明】

【0067】

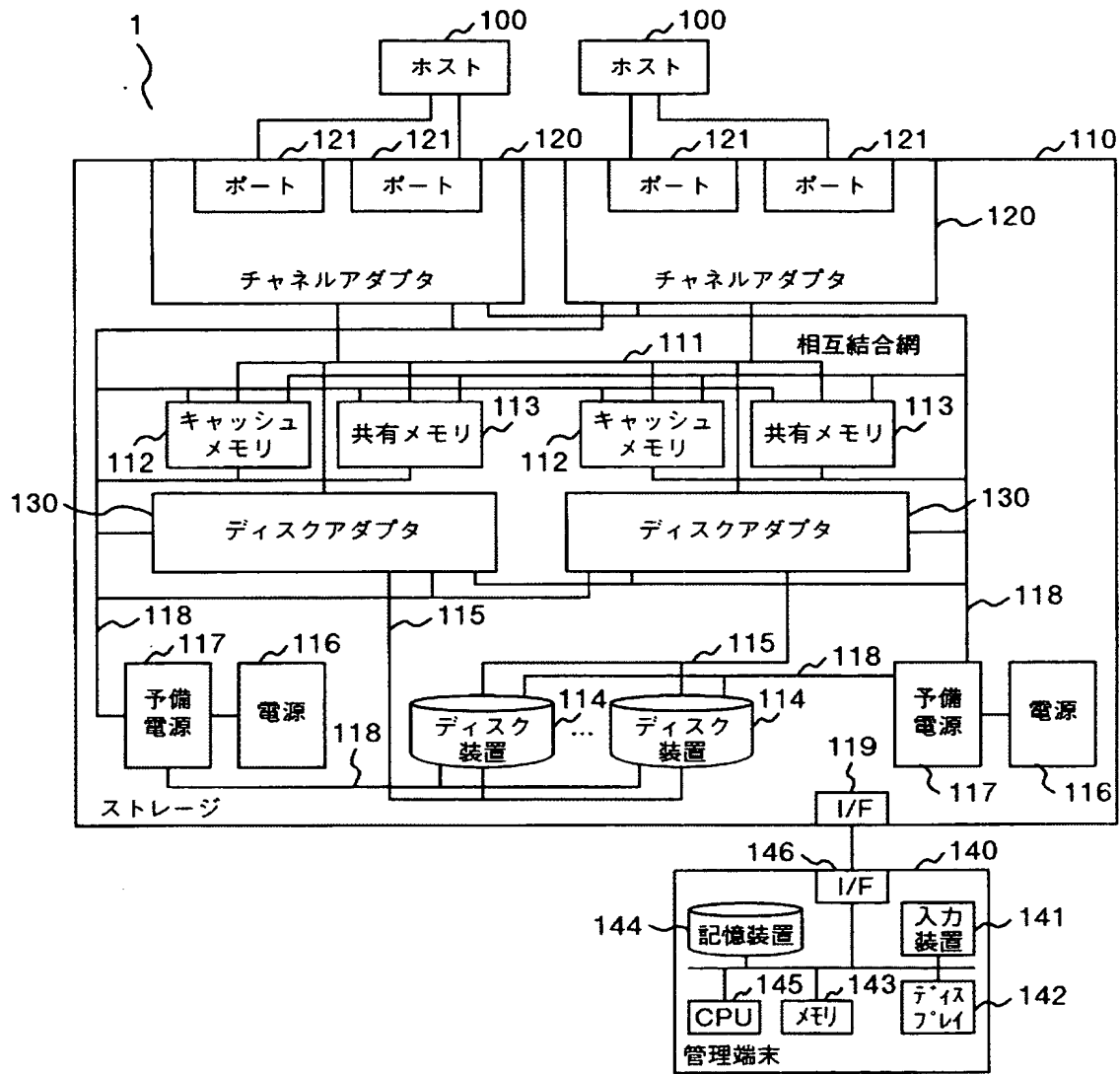
100…ホスト、110…ストレージ、111…相互結合網、112…キャッシュメモリ、113…共有メモリ、114…ディスク装置、115…接続線、116…電源、117…予備電源、118…電力線、120…チャンネルアダプタ、121…ポート、130…ディスクアダプタ、140…管理端末、141…入力装置、142…ディスプレイ、143…メモリ、144…記憶装置、145…CPU、146…I/F、

2…論理デバイス管理テーブル、3…L U パス管理テーブル、4…物理デバイス管理テーブル、5…スロット管理テーブル、6…セグメント管理テーブル、7…チャンネルジョブ管理テーブル、8…ディスクジョブ管理テーブル、9…ホスト管理テーブル、10…アクセスパターン管理テーブル、11…業務管理テーブル、12…スケジューリング管理テーブル、13…論理デバイス定義処理、14…L U パス定義処理、15…論理デバイス優先度定義処理、18…チャンネルアダプタポート処理（チャンネルアダプタ）、19…ジョブスケジューリング処理（チャンネルアダプタ）、20…リードジョブの処理（チャンネルアダプタ）、21…ライトジョブの処理（チャンネルアダプタ）、22…ジョブスケジューリング処理（ディスクアダプタ）、24…リードジョブの処理（ディスクアダプタ）、25…ライトジョブの処理（ディスクアダプタ）、16…ホスト優先度定義処理、17…業務優先度定義処理。

【書類名】 図面

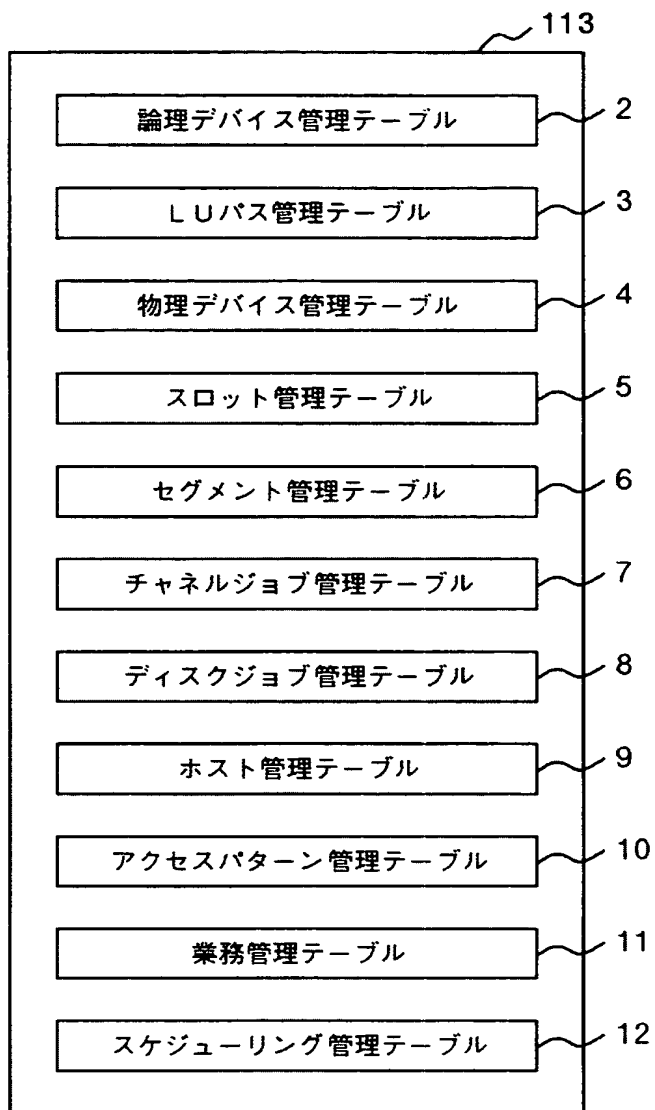
【図 1】

図 1



【図 2】

図 2



【図 3】

図 3

2

201	論理デバイス番号	1	2	3	4	5	6	7	..
202	サイズ	1GB	2GB	3GB	4GB	4GB	4GB	4GB	..
203	デバイス状態	オンライン	オンライン	オンライン	オフライン	オンライン	オンライン	オンライン	..
204	物理デバイス番号	1	2	3	4	5	6	7	..
205	接続ホスト番号	1	2	3	未定義	4	未定義	未定義	..
206	ポート番号 / ターゲットID / LUN	1/0/0	2/0/0	3/0/0	未定義	4/0/0	未定義	未定義	..
207	論理デバイス 優先度	1	5	3	3	1	3	3	..
208	拡張論理デバイス 情報	未定義	未定義	未定義	未定義	5, 6, 7	未定義	未定義	..
209	障害時データ 退避情報	未定義	未定義	未定義	未定義	未定義	未定義	未定義	..

【図 4】

図 4

3

301	302	303	304
ポート番号	ターゲットID / LUN	接続ホスト番号	対応論理デバイス 番号
1	0/0	1	1
2	0/0	2	2
:	:	:	:

【図 5】

図 5

4

401	物理デバイス番号	1	2	3
402	サイズ	72GB	72GB	72GB
403	対応論理デバイス 番号リスト	1, 2, 3, 4	5, 6, 7, 8	9, 10, 11, 12
404	デバイス状態	オンライン	オンライン	オンライン
405	RAID 構成	RAID5, 3D+1P	RAID5, 3D+1P	RAID1, 2D+2P
406	ディスク番号リスト	1, 2, 3, 4	5, 6, 7, 8	9, 10, 11, 12
407	ストライプサイズ	48KB	48KB	48KB
408	ディスク内サイズ	18GB	18GB	18GB
409	ディスク内開始 オフセット	0	0	0

【図 6】

図 6

5 501 スロット番号	502 セグメント番号 リスト	503 スロット属性	504 論理デバイス 番号	505 ホスト番号	506 ロック情報
1	1, 2, 3	クリーン	1	1	オン
2	4, 0, 6	ダーティ	2	2	オフ
3	7, 8, 9	ダーティ	3	3	オフ
4	0, 0, 0	フリー	4	4	オフ
:	:	:	:	:	:

【図 7】

図 7

6 601 セグメント番号	602 ブロック情報			
	1	2	3	4
1	有効	無効	有効	無効
2	無効	無効	有効	有効
3	有効	有効	無効	無効
:	:	:	:	:

【図 8】

図 8

7 701	ジョブ番号	1	2	3	..
702	処理	リード	リード	ライト	..
703	論理デバイス番号	1	2	3	..
704	転送開始位置	0	512	1024	..
705	転送長	512B	512B	512B	..
706	ホスト番号	1	2	3	..

【図 9】

図 9

8

801	ジョブ番号	1	2	3	..
802	処理	リード	リード	ライト	..
803	論理デバイス番号	1	2	3	..
804	転送開始位置	0	512	1024	..
805	転送長	512B	512B	512B	..
806	ホスト番号	1	2	3	..
807	チャネルジョブ番号	1	2	3	..

【図 10】

図 10

9

ホスト番号	ホスト名/WWN	ホスト優先度
1	hosta / WWNa	1
2	hostb / WWNb	2
3	hostc / WWNc	3
:	:	:

901 902 903

【図 11】

図 11

10

1001	論理デバイス番号	1	2	3	..
1002	リード回数	20	40	60	..
1003	ライト回数	30	50	70	..
1004	リードヒット回数	5	15	25	..
1005	ライトヒット回数	10	20	30	..
1006	シーケンシャル リード回数	1	3	5	..
1007	シーケンシャル ライト回数	2	4	6	..
1008	ダーティ量 管理情報	0.1G B.O.	0.1 GB,	0.1 GB	..

【図 12】

図 12

11

1101

1102

1103

1104

業務番号	論理デバイス番号	ホスト番号	業務優先度
1	0	0	1
2	0	1	2
3	1	0	3
:	:	:	:

【図 13】

図 13

12

1201

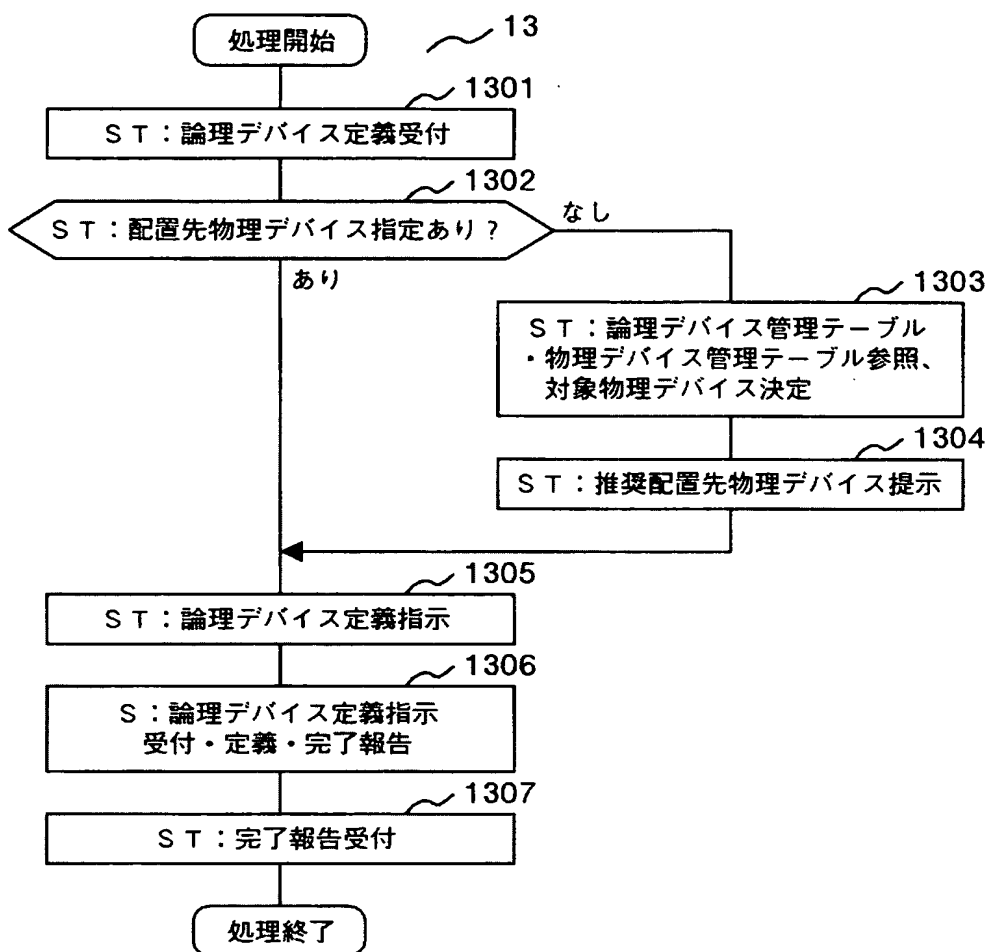
1202

1203

論理デバイス 優先度	ホスト優先度	業務優先度
0.5	0.3	0.2

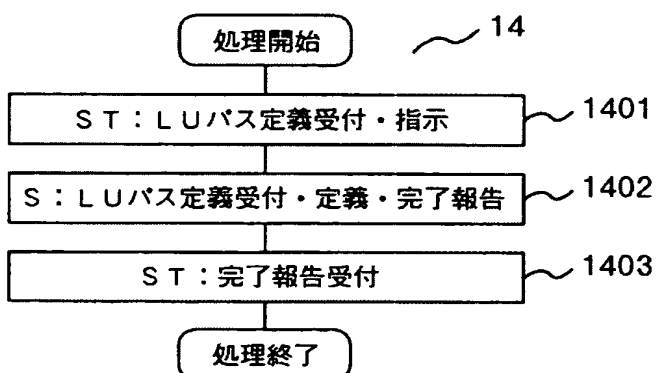
【図 14】

図 14



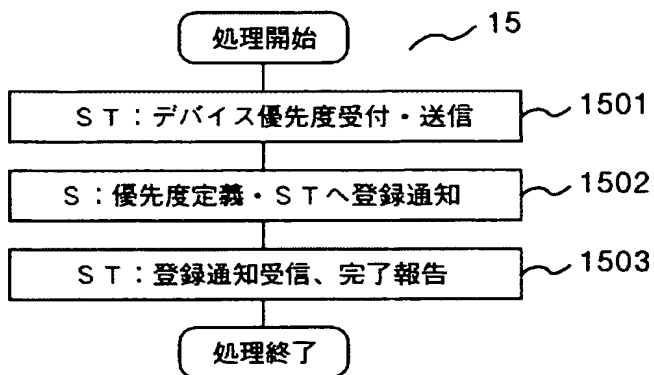
【図 15】

図 15



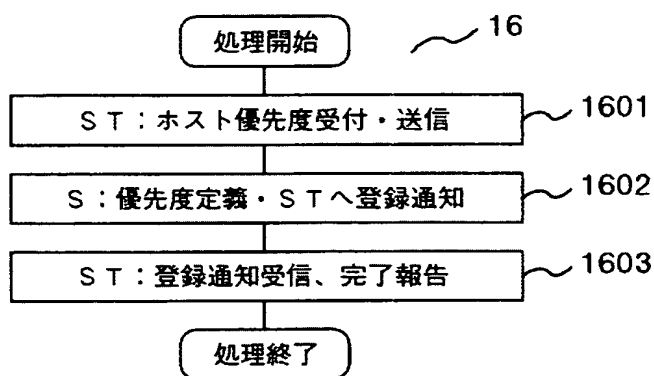
【図 16】

図 16



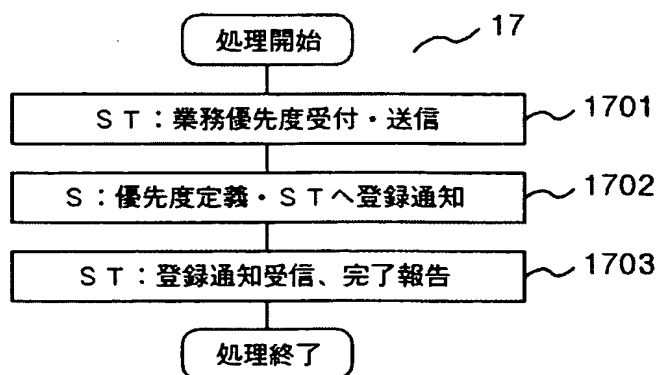
【図 17】

図 17



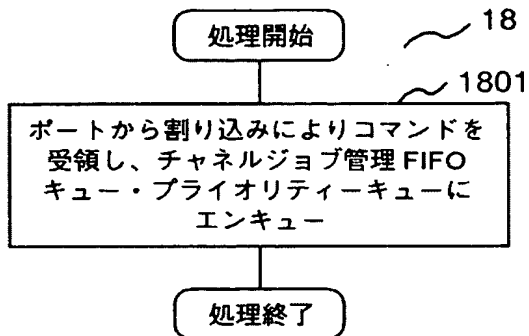
【図 18】

図 18



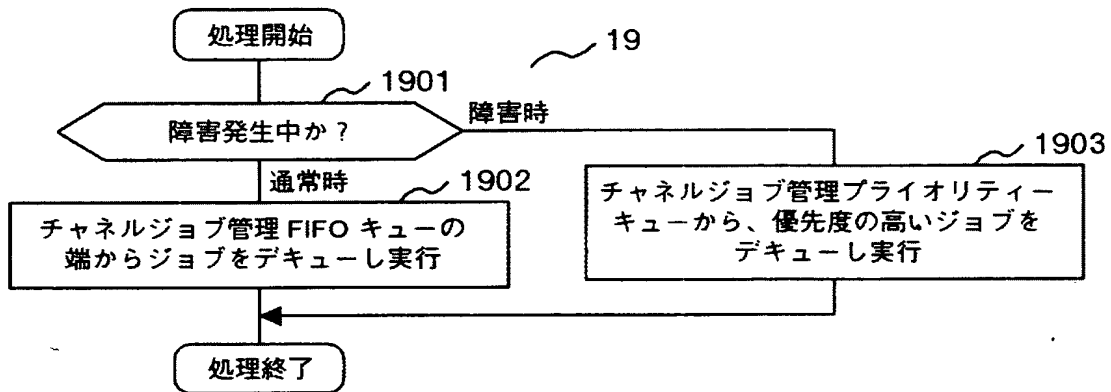
【図 19】

図 19



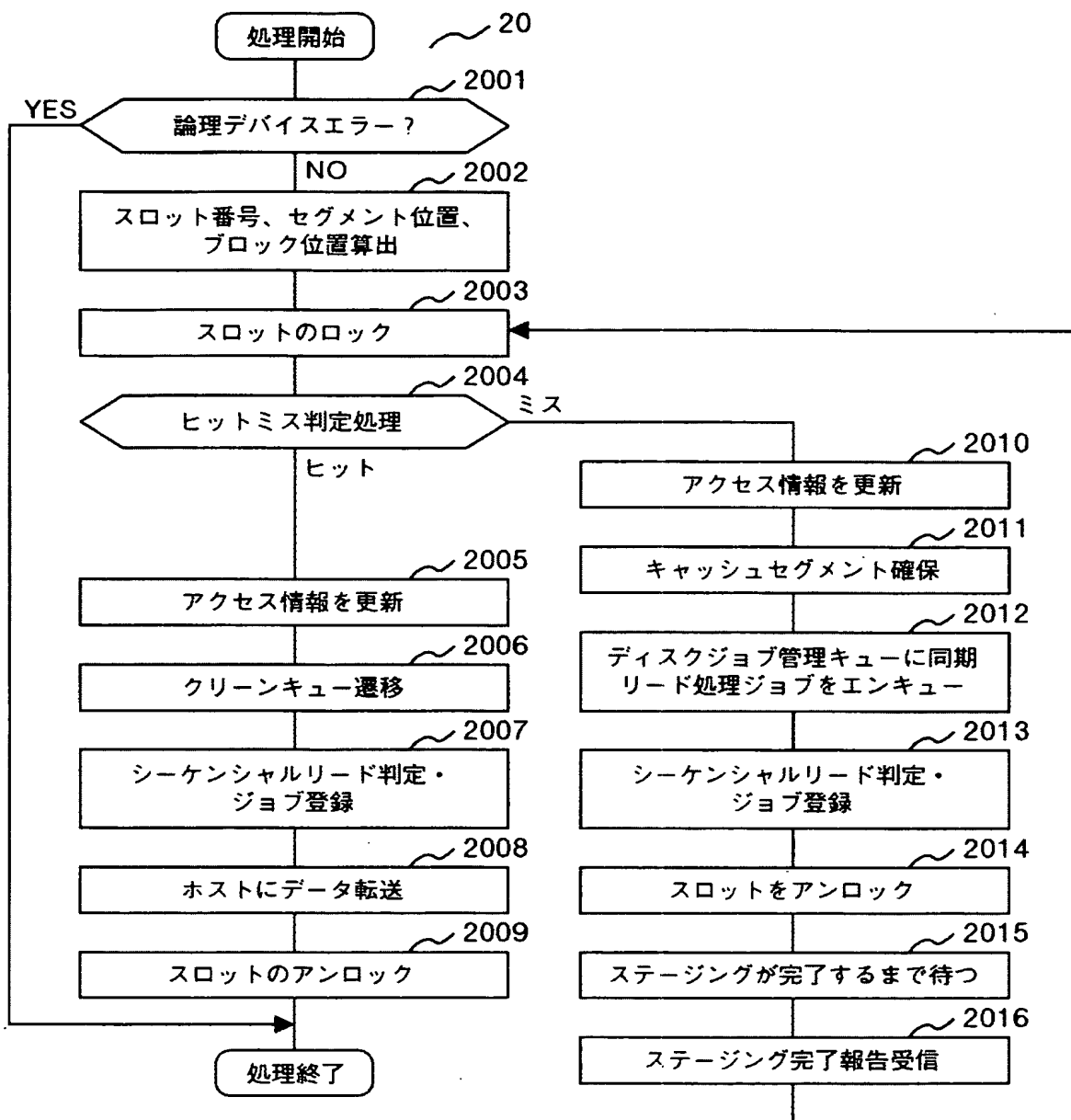
【図 20】

図 20



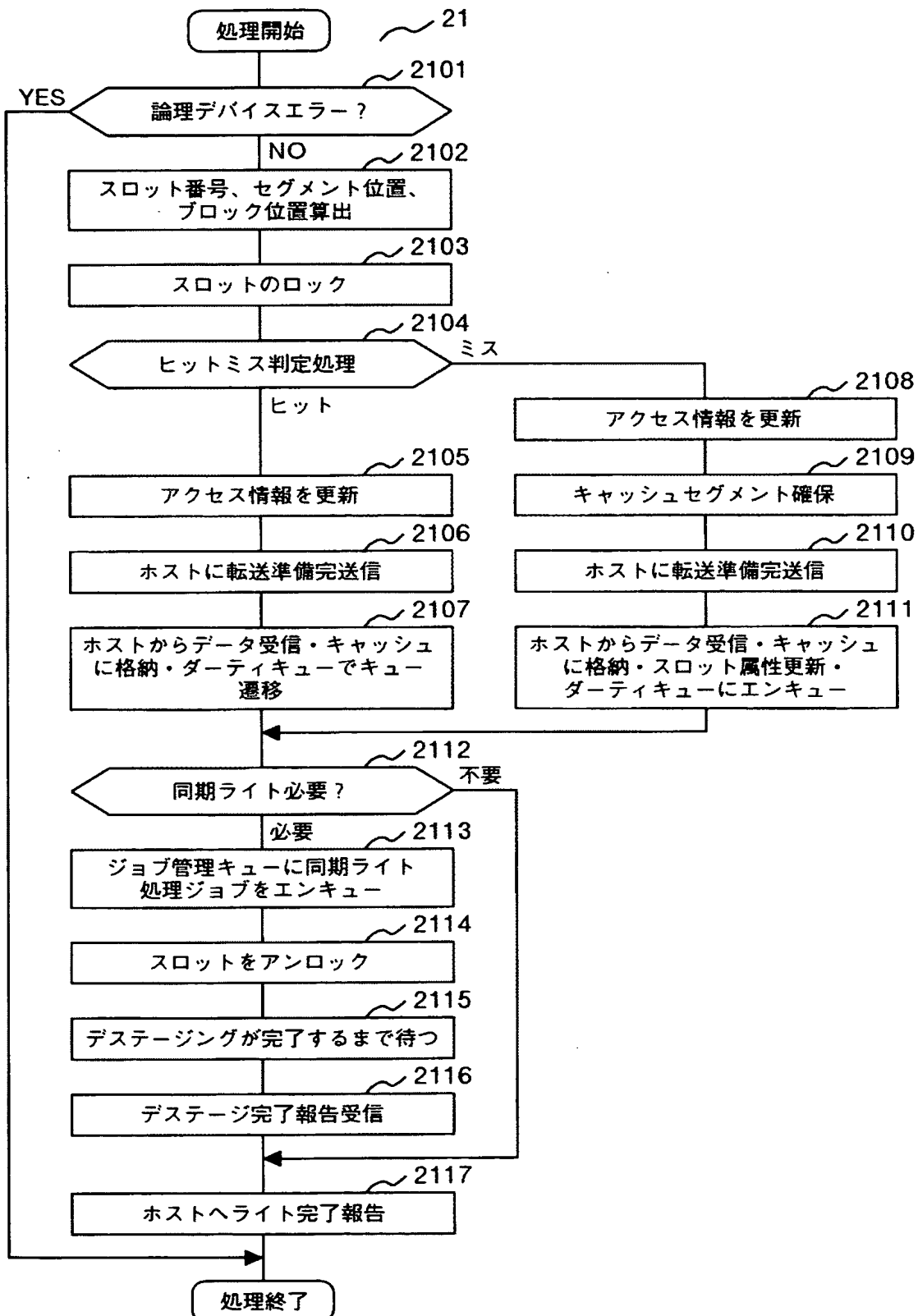
【図 21】

図 21



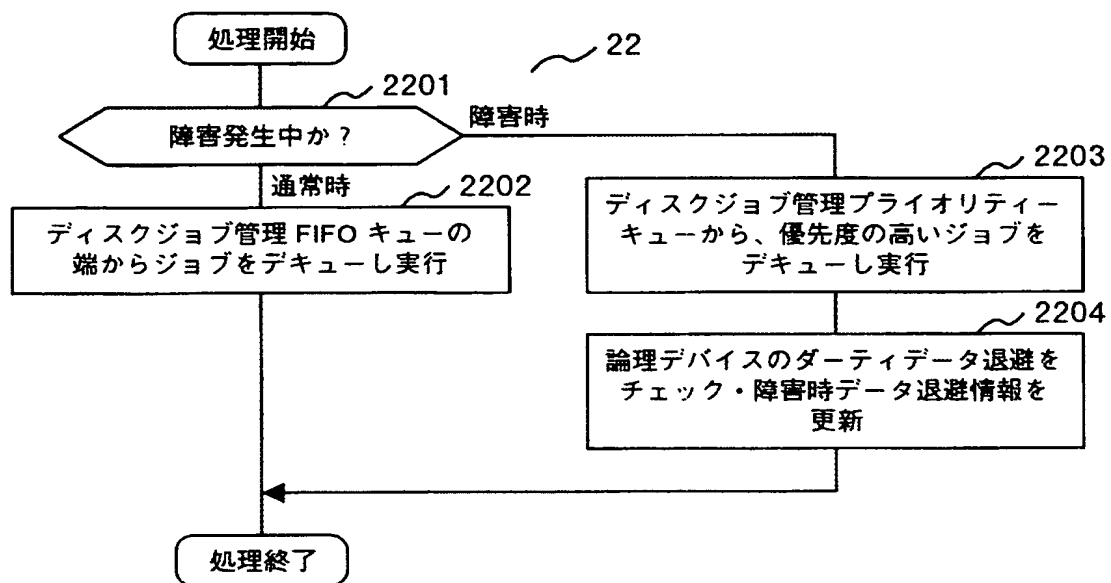
【図 22】

図 22



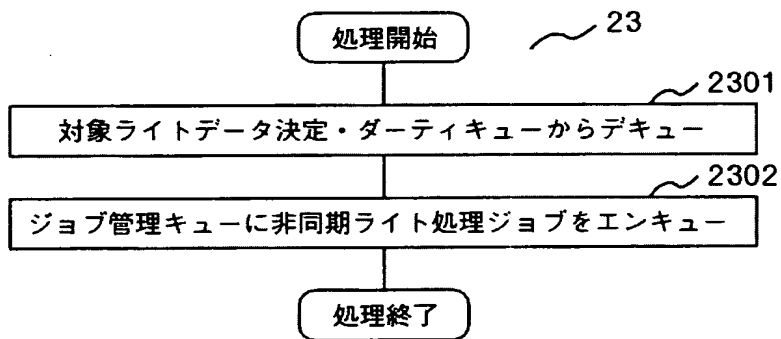
【図 23】

図 23



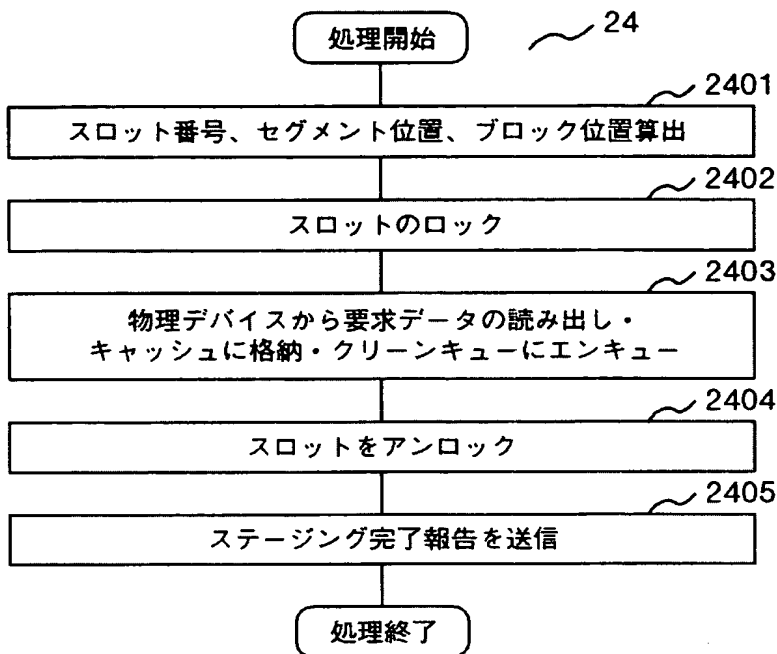
【図 24】

図 24



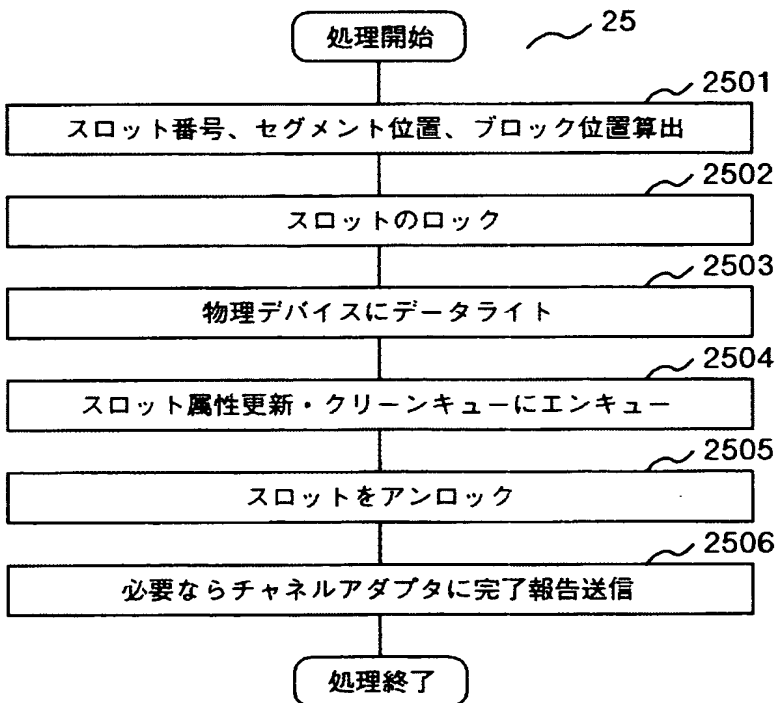
【図 25】

図 25



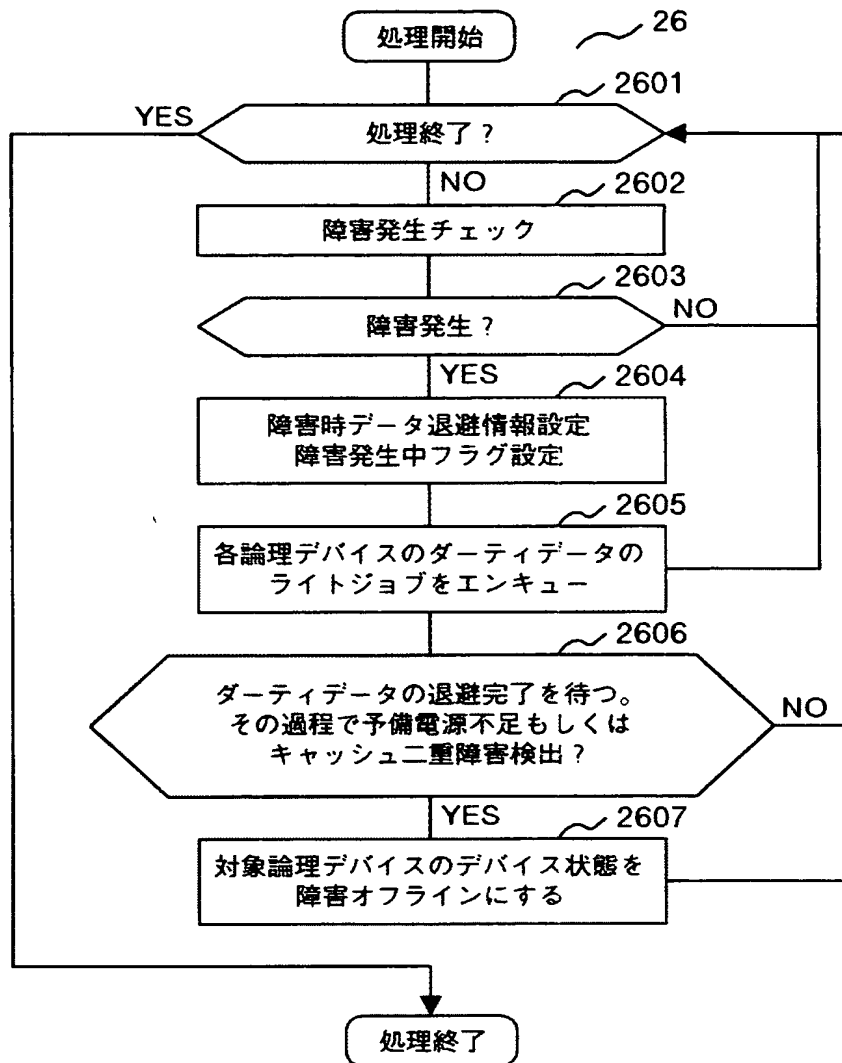
【図 26】

図 26



【図 27】

図 27



【書類名】 要約書**【要約】****【課題】**

キャッシュメモリとディスク装置を有するストレージにおいて、障害発生時に、優先度の高い重要なデータの消失を回避し、ストレージに対する入出力処理全体として低下し、優先度の高い業務に対して性能影響が生じるのを極力回避する。

【解決手段】

ホストに提供するデバイスの単位である論理デバイスを複数のディスク装置から構成される複数の物理デバイスに論理デバイスに設定された優先度を考慮して分散して配置することにより、障害発生時に優先度の基づいて速やかにダーティデータをディスク装置に反映する。また、障害発生時に重要な業務に関するジョブを優先的に処理することにより、ホスト処理性能の低下を軽減する。

【選択図】 図 1

特願 2 0 0 3 - 3 9 0 2 3 9

出 願 人 履 歴 情 報

識別番号

[0 0 0 0 0 5 1 0 8]

1 . 変 更 年 月 日

1 9 9 0 年 8 月 3 1 日

[変 更 理 由]

新 規 登 録

住 所

東 京 都 千 代 田 区 神 田 駿 河 台 4 丁 目 6 番 地

氏 名

株 式 会 社 日 立 製 作 所